The EMBO Journal (2011) 30, 5010–5020 | © 2011 European Molecular Biology Organization | All Rights Reserved 0261-4189/11
www.embojournal.org

THE EMBO JOURNAL

Structure of the Shq1–Cbf5–Nop10–Gar1 complex and implications for H/ACA RNP biogenesis and dyskeratosis congenita

Shuang Li[1,2,3], Jingqi Duan[2,3], Dandan Li[2], Shoucai Ma[2] and Keqiong Ye[2,*]

[1]Department of Biochemistry and Molecular Biology, College of Life Sciences, Beijing Normal University, Beijing, China and [2]National Institute of Biological Sciences, Beijing, China

Shq1 is a conserved protein required for the biogenesis of eukaryotic H/ACA ribonucleoproteins (RNPs), including human telomerase. We report the structure of the Shq1-specific domain alone and in complex with H/ACA RNP proteins Cbf5, Nop10 and Gar1. The Shq1-specific domain adopts a novel helical fold and primarily contacts the PUA domain and the otherwise disordered C-terminal extension (CTE) of Cbf5. The structure shows that dyskeratosis congenita mutations found in the CTE of human Cbf5 likely interfere with Shq1 binding. However, most mutations in the PUA domain are not located at the Shq1-binding surface and also have little effect on the yeast Cbf5–Shq1 interaction. Shq1 binds Cbf5 independently of the H/ACA RNP proteins Nop10, Gar1 and Nhp2 and the assembly factor Naf1, but shares an overlapping binding surface with H/ACA RNA. Shq1 point mutations that disrupt Cbf5 interaction suppress yeast growth particularly at elevated temperatures. Our results suggest that Shq1 functions as an assembly chaperone that protects the Cbf5 protein complexes from non-specific RNA binding and aggregation before assembly of H/ACA RNA.

The EMBO Journal (2011) **30,** 5010–5020. doi:10.1038/emboj.2011.427; Published online 25 November 2011
Subject Categories: RNA; structural biology
Keywords: assembly chaperone; crystal structure; dyskeratosis congenita; H/ACA RNP biogenesis

Introduction

Box H/ACA RNAs constitute a large family of non-coding RNAs conserved in archaea and eukaryotes (Meier, 2005; Matera *et al*, 2007; Kiss *et al*, 2010). Each distinct H/ACA RNA assembles with a common set of four proteins, Cbf5 (NAP57 in rodents and dyskerin in humans), Nop10, Nhp2 (L7Ae in archaea) and Gar1 to form an H/ACA ribonucleoprotein (RNP) particle. The function of most H/ACA RNPs is to convert uridine into pseudouridine at numerous specific sites within ribosomal RNAs (rRNAs) and spliceosomal small nuclear RNAs (snRNAs), and the substrate specificity is determined by the H/ACA RNA. However, the snR30/U17 H/ACA RNP and human telomerase are involved in precursor rRNA processing and telomere formation, respectively. Human telomerase contains the reverse transcriptase hTERT, the telomerase RNA hTR as well as the four H/ACA proteins that associate with an H/ACA domain in hTR, which is required for telomerase *in-vivo* assembly, localization and activity (Mitchell *et al*, 1999).

Dyskeratosis congenita (DC) is a rare genetic disease characterized by a classic triad of nail dystrophy, abnormal skin pigmentation and oral leukoplakia as well as by bone marrow failure, pulmonary fibrosis, cancer and other complications (Walne and Dokal, 2009; Bessler *et al*, 2010). DC is considered as a telomere deficiency disorder, because DC patients show very short telomere in highly proliferating tissues and the causal mutations found so far are all located in genes that control telomere homeostasis, including the telomerase components dyskerin, hTR, hTERT, Nop10 and Nhp2, the telomere protecting protein TIN2 and the telomerase trafficking factor TCAB1 (Zhong *et al*, 2011). Mutations in dyskerin cause the most frequent X-linked form of DC (Heiss *et al*, 1998). DC mutations have been shown to interfere with telomerase stability, assembly, activity and localization (Mochizuki *et al*, 2004; Trahan and Dragon, 2009; Robart and Collins, 2010; Trahan *et al*, 2010; Batista *et al*, 2011).

In eukaryotes, mature H/ACA RNPs that modify rRNAs are localized in the nucleolus, the place of ribosome assembly, and those that modify snRNAs are found in the Cajal bodies (Darzacq *et al*, 2002). According to the different localization, H/ACA RNA/RNPs are classified as small nucleolar (sno) and small Cajal body-specific (sca) RNA/RNPs. H/ACA scaRNAs, including hTR, are directed to the Cajal bodies by a CAB motif in the apical loop (Richard *et al*, 2003; Jady *et al*, 2004).

An H/ACA RNA and the four core proteins can assemble spontaneously *in vitro* into an active enzyme in both archaeal and eukaryotic systems (Baker *et al*, 2005; Charpentier *et al*, 2005; Li *et al*, 2011). However, assembly of eukaryotic H/ACA RNPs is highly complicated *in vivo* and requires several general as well as H/ACA-specific assembly factors (Kiss *et al*, 2010). The chaperone heat-shock protein 90 (Hsp90) and its associated proteins have been shown to be involved in assembly of H/ACA, C/D and other RNPs that all contain a L7Ae-related protein (Boulon *et al*, 2008; Venteicher *et al*, 2008; Zhao *et al*, 2008). Two conserved proteins, Naf1 and Shq1, are specifically required for H/ACA RNP formation in yeast *Saccharomyces cerevisiae* (Dez *et al*, 2002; Fatica *et al*, 2002; Yang *et al*, 2002) and in mammalian cells (Darzacq *et al*, 2006; Hoareau-Aveilla *et al*, 2006; Grozdanov *et al*, 2009b). Naf1 and Shq1 are localized in the nucleoplasm and excluded from nucleoli and Cajal bodies, where mature H/ACA RNPs reside (Dez *et al*, 2002; Fatica *et al*, 2002; Yang *et al*, 2002; Darzacq *et al*, 2006; Grozdanov *et al*, 2009b).

*Corresponding author. National Institute of Biological Sciences, Beijing 102206, China. Tel.: +86 10 8072 6688/ext. 8550; Fax: +86 10 8072 8592; E-mail: yekeqiong@nibs.ac.cn
[3]These authors contributed equally to this work

Received: 1 August 2011; accepted: 7 November 2011; published online: 25 November 2011

H/ACA RNP proteins Cbf5, Nhp2 and Nop10 and assembly factor Naf1 are recruited to the site of H/ACA RNA genes, suggesting that the four proteins assemble with the nascent H/ACA RNAs into precursor RNPs (pre-RNPs) (Ballarino *et al*, 2005; Yang *et al*, 2005; Darzacq *et al*, 2006). Naf1 shares a homologous core domain with Gar1 (Leulliot *et al*, 2007) and associates with Cbf5 (Darzacq *et al*, 2006). We recently showed that the H/ACA pre-RNP assembled with Naf1 possesses basic pseudouridylation activity but is incapable of substrate turnover and replacement of Naf1 by Gar1 yields a fully active enzyme (Li *et al*, 2011).

Shq1 appears to be an early assembly factor that, unlike Naf1, is not associated with the H/ACA RNA genes in yeast and mammalian cells (Yang *et al*, 2005; Grozdanov *et al*, 2009b). Shq1 is composed of an N-terminal CS (CHORD-containing proteins and Sgt1) domain and a C-terminal Shq1-specific domain (SSD). The structure of the CS domain has been determined by crystallography and NMR (Godin *et al*, 2009; Singh *et al*, 2009). The CS domain adopts a β-sandwich structure similar to that of Hsp20 superfamily proteins, which are co-chaperones of Hsp90. However, no interaction has been found between the CS domain of Shq1 and yeast Hsp90 (Godin *et al*, 2009; Singh *et al*, 2009). The SSD interacts with the PUA domain of Cbf5 and some of DC mutations in the PUA domain were shown to modulate Shq1 binding (Grozdanov *et al*, 2009a). The molecular function of Shq1 remains unclear.

In this study, we analyse the Shq1–Cbf5 interaction by using X-ray crystallography, biochemical approaches, mutagenesis and yeast experiments. Our results reveal the structural details of Shq1–Cbf5 recognition and provide novel insights into the function of Shq1 in H/ACA RNP biogenesis and the effect of DC mutations in dyskerin.

Results

Shq1 binds Cbf5 independently of other Cbf5-interacting proteins

The isolated Cbf5 protein has poor solubility, impeding analysis of its interaction with Shq1. By co-expression and co-purification of yeast Cbf5 with Nop10 or with both Nop10 and Gar1, we recently obtained soluble Cbf5 complexes of various compositions that are functional in RNP assembly and pseudouridylation (Li *et al*, 2011). The structure of the yeast Cbf5–Nop10–Gar1 (CNG) ternary complex has also been determined (Li *et al*, 2011). These progresses allowed us to examine the interaction of Shq1 with Cbf5 complexes in detail (Figure 1A). Yeast Cbf5 possesses a core domain (residues 1–394) and a C-terminal KKE/D repeat tail and yeast Gar1 has a core domain flanked by two genetically dispensable Gly/Arg-rich (GAR) domains (Girard *et al*, 1994; Pogacic *et al*, 2000). Unless otherwise indicated, we used the core domain of Cbf5 and Gar1 and the Gar1-like core domain of Naf1 (residues 107–280) for biochemical and structural studies (Leulliot *et al*, 2007).

We incubated purified full-length Shq1 with the Cbf5 complexes and used Ni beads to pull down His-tagged Cbf5. The Cbf5–Nop10 complex retained a stoichiometric amount of Shq1, whereas the Ni beads alone retained only a background amount of Shq1, indicating a specific interaction between Shq1 and the Cbf5–Nop10 complex. This result is somewhat surprising as a previous study with *in-vitro* translated rat proteins showed that Shq1 only binds NAP57 (rat Cbf5) alone but not NAP57 in complex with Nop10, Nhp2 and Naf1 (Grozdanov *et al*, 2009b). As expression of Cbf5 alone produced protein aggregates that cannot be purified, we were not able to examine whether Cbf5 has a higher binding affinity for Shq1 than the Cbf5–Nop10 complex. Nevertheless, our results indicate that Cbf5 can bind simultaneously to Shq1 and Nop10.

We further examined whether the other H/ACA RNP proteins Nhp2 and Gar1 and the assembly factor Naf1 affect the Shq1 interaction. The incorporation of these proteins in various combinations into the Cbf5–Nop10 complex also had no effect on Shq1 retention (Figure 1A). Moreover, the compositions of the retained Cbf5 complexes were not changed in all cases before and after Shq1 binding, indicating that Shq1 does not disrupt the pre-assembled Cbf5 complex structures. Shq1 was similarly bound by the Cbf5 complexes assembled with full-length Cbf5 and full-length Gar1 (partially degraded), indicating that the C-terminal KEE/D tail of Cbf5 and the GAR domains of Gar1 do not interfere with the Shq1 association (Figure 1B).

Isothermal titration calorimetry (ITC) measurements showed that Shq1 binds tightly to Cbf5–Nop10 with an apparent dissociation constant (K_d) of 45 ± 27 nM (Figure 1C). The association of Gar1 and Nhp2 did not affect the binding affinity within experimental error. We can conclude that Shq1 contacts Cbf5 independently of the other Cbf5-interacting proteins *in vitro*.

Crystal structure of the Shq1-specific domain

The C-terminal SSD of Shq1 has been shown to interact with Cbf5 (Godin *et al*, 2009; Grozdanov *et al*, 2009b). We confirmed this interaction using our recombinant yeast proteins by pull-down assay and ITC (unpublished observation). To further understand the function of the SSD, we determined its crystal structure by selenium phasing. The structure was refined to 1.6 Å resolution with an R_{work}/R_{free} of 0.168/0.198 (Table 1; Supplementary Figure S1). The structure shows that the SSD encompasses residues 164 to close to the C-terminus of Shq1 and adopts a unique globular fold with dimensions of $\sim 73 \times 66 \times 41$ Å (Figure 2A–D). The structure is mostly helical and consists of 17 α-helices (α1–α17) and 2 short β-strands (β1–β2). Helix α8 is located at the centre of the structure and is wrapped by two clockwise circles of α-helices. Helices α1–α7 comprise the outer ring, whereas helices α9–α16 form the inner ring. Strands β1 and β2, which form a β-hairpin, and helix α17 form a V-shaped structure packed at the outside. Two loops connecting helices α10 and α11 (residues 346–362) and strands β1 and β2 (residues 472–482) were not visible in the crystal and are likely disordered.

The fold of the SSD is unprecedented. Structural homologues (Z-score <3.6) identified by the Dali server show limited similarity only to helices α8–α12 of the SSD structure (Holm and Sander, 1993). This region adopts a 4-helix bundle, a common motif embedded in many proteins with diverse functions, including cyclin and retinoblastoma tumour suppressor. The implication of this limited structure homology is not clear. The structure also contains an unusual buried charge cluster that is composed of highly conserved arginine and acidic residues and mediates the packing of helices α2 and α11 (Supplementary Figure S2; Figure 2D).

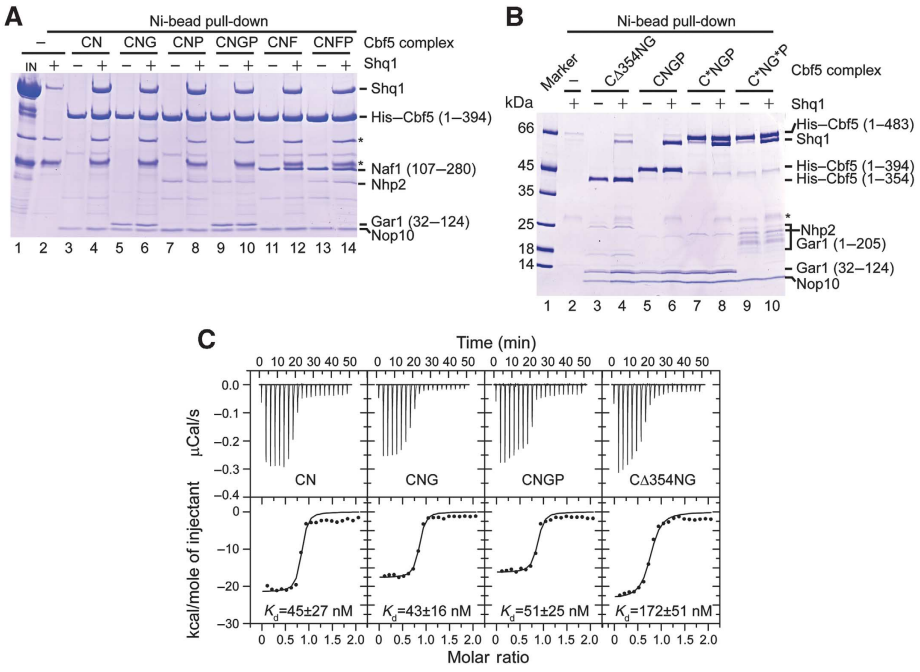

Figure 1 Interaction between Shq1 and the Cbf5 complexes. (**A**) Shq1 binds Cbf5 independently of other proteins. The Cbf5 complexes are denoted by concatenated abbreviations of the constituent proteins: C for Cbf5 (1–394), CΔ354 for Cbf5 (1–354), N for Nop10, G for Gar1 (32–124), P for Nhp2 and F for Naf1 (107–280). The Cbf5 complexes were incubated with Shq1 and Ni beads in 50 mM sodium phosphate (pH 7.6) and 0.4 M KCl for 2 h at 4°C. Among these proteins, only Cbf5 contained a His tag. Half of the bound protein was analysed on Coomassie blue-stained SDS–PAGE gels. Lane 1: 100% of the Shq1 input. Lane 2: non-specifically retained Shq1. The migration position and amino-acid boundaries of each protein are labelled. The asterisks indicate unidentified proteins from the Shq1 sample that bound strongly to Ni beads. (**B**) Shq1 pull-down assay with Cbf5 complexes containing full-length Cbf5 (C*), Cbf5 (1–354) or full-length Gar1 (G*). Full-length Gar1 was partially degraded. (**C**) ITC analysis of Shq1–Cbf5 interaction. Shq1 was titrated into the indicated Cbf5 complexes. The upper panels show the heat change upon injection, and the lower panels show the integrated heat curves that were fit using a one-set-of-sites model. The apparent dissociation constants are indicated.

Crystal structure of the Shq1-specific domain in complex with Cbf5, Nop10 and Gar1

To reveal the molecular detail of the Shq1–Cbf5 interaction, we determined the crystal structure of the SSD in complex with Cbf5, Nop10 and Gar1 at 3.06 Å resolution. The structure was solved by molecular replacement and refined to an R_{work}/R_{free} of 0.218/0.275 (Table 1).

The quaternary complex structure shows that Shq1 exclusively contacts Cbf5 at the PUA domain, the C-terminal extension (CTE) and a small part of the catalytic domain (Figures 3A and 4). The Shq1–Cbf5 complex buries 1803 Å^2 of solvent accessible area in Cbf5, with 612 Å^2 contributed by the PUA domain, 1082 Å^2 by the CTE and 109 Å^2 by the catalytic domain.

In the Shq1 complex structure, the Shq1-binding interface on Cbf5 is well separated from that of Nop10 and Gar1. The assembly factor Naf1 most likely binds Cbf5 at the same site as Gar1, and Nhp2 probably binds Nop10 (Li and Ye, 2006; Li *et al*, 2011). These structural observations and predictions indicate that Shq1 would not interfere with the interaction of Cbf5 with these proteins, in agreement with our biochemical results (Figure 1). The PUA domain is also the anchoring site for the lower stem and the ACA motif of H/ACA RNA (Li and Ye, 2006). Alignment of the *Pyrococcus furiosus* (Pf) H/ACA RNP structure shows that Shq1 and the H/ACA RNA share a largely overlapping interface on the PUA domain, indicating that they interact with Cbf5 in a mutually exclusive manner (Figure 3B).

Structural change upon Shq1–Cbf5 association

The structures of SSD alone and in the Cbf5 complex can be superimposed over the entire structure with an r.m.s.d. of 0.303 Å over 298 Cα pairs (Supplementary Figure S3). This indicates that Shq1 provides a rather rigid interface for Cbf5 interaction. One exception is the C-terminal β-hairpin of Shq1, the apical portion of which is disordered in the free state but is largely structured upon Cbf5 binding.

Compared with the CNG structure (Li *et al*, 2011), the Cbf5 structure undergoes several changes upon Shq1 association. Most dramatically, the Cbf5 CTE, which was flexible and degraded in the CNG structure, becomes largely structured in the Shq1 complex. The PUA domain rotates 8° relative to the catalytic domain owing to the plastic connection between them, as previously seen in the archaeal Cbf5 structure (Duan *et al*, 2009). At the centre of the Shq1-binding interface, local conformational changes were observed in the loop connecting helix α8 and strand β3′. However, this region in the CNG structure is involved in crystal packing and its conformation in the Shq1 complex better resembles that in archaeal Cbf5. The N-terminal residues 1–17 of Cbf5 were not visible in the Shq1 complex structure, probably because of degradation. Loop L1 formed by residues 5–11 was only occasionally observed with high temperature factors in the CNG complex structure (Li *et al*, 2011).

Shq1–Cbf5 interaction

Shq1 contacts Cbf5 mainly through charge complementarity and hydrophobic contacts. At the centre of the interface, the

Table 1 Data collection and refinement statistics

Crystal form	Native SSD	Se SSD	SSD–Cbf5–Nop10–Gar1
Data collection			
Space group	P1	P1	$P4_32_12$
Cell dimensions			
a, b, c (Å)	39.1, 50.1, 54.0	39.2, 50.1, 54.0	101.0, 101.0, 272.3
α, β, γ (deg)	102.2, 93.8, 98.0	102.3, 94.0, 97.5	90, 90, 90
Wavelength (Å)	0.9792	0.9793	0.9789
X-ray source	SSRF BL17U1	SSRF BL17U1	SSRF BL17U1
Resolution range (Å)[a]	50–1.6 (1.63–1.60)	50–2.1 (2.14–2.1)	50–3.06 (3.11–3.06)
Unique reflections	50 317	20594	27 487
Redundancy	3.9 (4.0)	3.6 (3.6)	14.1 (14.7)
I/σ	42.5 (14.1)	25.1 (10.2)	58.4 (5.6)
Completeness (%)	96.8 (95.3)	89.4 (96.7)	99.9 (100)
R_{merge}[b]	0.057 (0.132)	0.091 (0.173)	0.067 (0.684)
Structure refinement			
Resolution range (Å)	20–1.6 (1.64–1.6)		20–3.06 (3.14–3.06)
No. of reflections	47 740		25 934
No. of atoms	2929		6627
Mean B factor (Å^2)	13.6		100.6
R_{work}[c]	0.168 (0.173)		0.218 (0.279)
R_{free}[d]	0.198 (0.201)		0.275 (0.332)
R.m.s.d. bond length (Å)	0.010		0.009
R.m.s.d. bond angles (deg)	1.212		1.210
RAMPAGE statistics[e]			
Favoured (%)	98.4		92.1
Allowed (%)	1.3		7.0
Outlier (%)	0.3		0.9

[a]The values for the data in the highest resolution shell are shown in parentheses.
[b]$R_{merge} = \Sigma|I_i - I_m|/\Sigma I_i$, where I_i is the intensity of the measured reflection and I_m is the mean intensity of all symmetry related reflections.
[c]$R_{work} = \Sigma_h|F_o - F_c|/\Sigma_h F_o$, where F_o and F_c are the observed and calculated structure factor amplitudes of reflection h.
[d]R_{free} is the same as R_{work}, but calculated on 5% reflections not used in refinement.
[e]Analysed by RAMPAGE including general residues, glycine, proline and pre-proline.

N-terminal part of helix α9 and loops α8–α9 and α11–α12 of Shq1 binds helix α8, strands β3′ and β17, and the loop C-terminal to β17 of Cbf5 (Figure 4C). Specifically, the planar guanidinium group of Cbf5 Arg343 stacks over the indole ring of Shq1 Trp326 and forms a salt bridge with the side chain of Shq1 Glu323. The aromatic ring of Cbf5 Try281 packs against the hydrophobic side chain of Shq1 Leu381 and makes a hydrogen bond with the side chain amino of mostly buried Shq1 Arg383, which is also within the electrostatic contact distance of Cbf5 Asp344. These Cbf5-interacting residues (Trp326, Glu323, Leu381 and Arg383) are highly conserved in Shq1 and form a small surface patch, in contrast to a general lack of conservation in the other surface areas of the Shq1 structure (Figures 2D and 4D). Consistent with their structural role in Cbf5 binding and the high degree of conservation, the replacement of the hydrophobic Trp326 residue with a charged aspartate and the substitution of Arg383 with a negatively charged glutamate significantly reduced the Shq1–Cbf5 interaction in ITC experiments (Figure 4F). These mutant proteins showed the same elution profile in sizing columns as wild-type Shq1 and some residual binding to Cbf5 in ITC, indicating that they remain correctly folded.

At the periphery of the interface, the loop C-terminal to helix α1 and residue Asp421 in loop α13–α14 of Shq1 contact a small area of the Cbf5 catalytic domain, which is around helix α2 and adjacent to the PUA domain (Figure 4C). Shq1 Asp421 makes an electrostatic contact with Cbf5 Lys114. Although the interacting pair is conserved in the respective proteins, mutation of Asp421 to lysine did not apparently affect the binding affinity (Figure 4F), suggesting that this interaction makes a minor energetic contribution. At the other side of the interface, helix α4 of Shq1, which however is not conserved in the human protein, contacts helix α8 and loop β3′–β2′ of Cbf5 (Figure 4A and C).

The CTE was not observed in the CNG structure because of flexibility and degradation (Li *et al*, 2011). In the Shq1 complex structure, the CTE to residue 378 forms an extended structure over the surface of Shq1 (Figure 4D and E). The CTE contains α-helix α3′ (residues 355–364) flanked by two loops. The loop connecting helix α3′ to the PUA domain contacts helices α3 and α17 of Shq1 mainly via the hydrophobic residues Trp350 and Leu352 (Figure 4D). The α3′ helix of Cbf5 packs perpendicularly against the inner walls of the V-shaped structure formed by the C-terminal β1–β2 hairpin and helix α17 of Shq1, and the C-terminal loop of Cbf5 hooks at the edge of the Shq1 structure. In this region, the Shq1–Cbf5 interaction is mediated by numerous electrostatic interactions mainly between basic residues in Cbf5 and acidic residues in Shq1 (Figure 4A and D). In addition, Cbf5 Met362, Val374 and Leu368 and Shq1 Leu494 form a small hydrophobic core at the molecular interface (Figure 4D). Consistent with the structural observation, truncation of the C-terminal region of Cbf5 after residue 354 (Cbf5Δ354), which removes helix α3′ and the following loop, reduces the Shq1-binding affinity by ~4-fold (Figure 1B and C). The structure shows that the conserved negatively charged residue Glu219 of Shq1 caps the positively charged N-terminus of Cbf5 helix α3′ and likely neutralizes the helix dipole (Figure 4D). However, mutation E219K has little effect on Cbf5 binding (Figure 4F).

Shq1 mutants in yeast

To investigate the functional importance of the Shq1–Cbf5 interaction, we made a *Shq1* knockout haploid strain and complemented it with *Shq1* mutants (Figure 5A). Deletion of

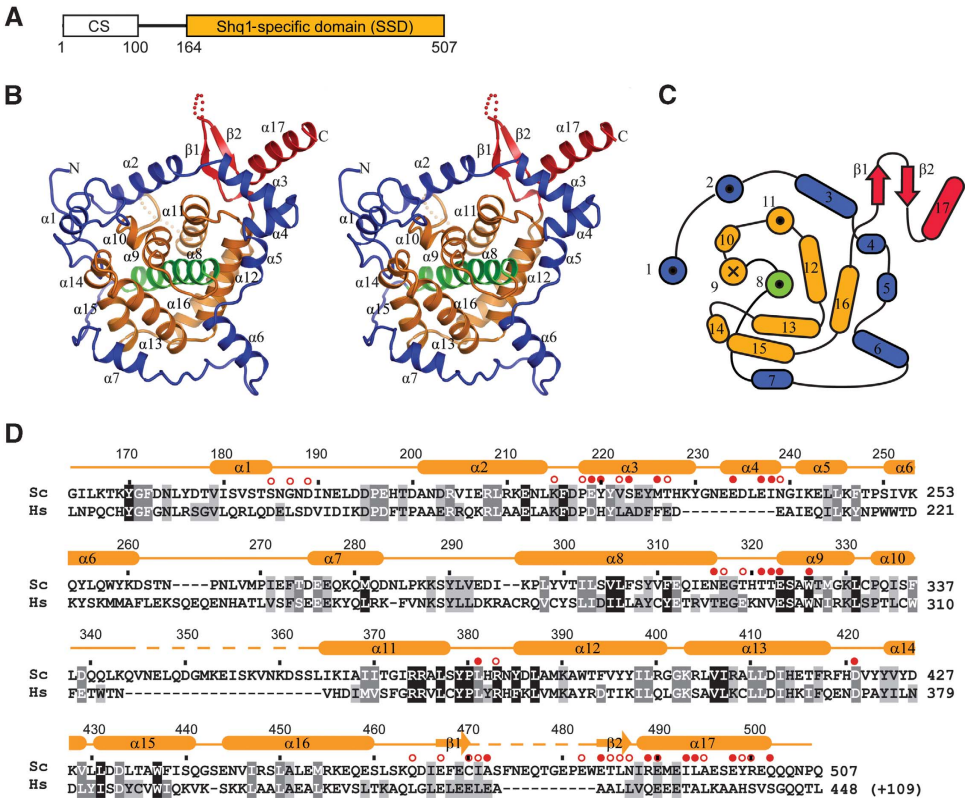

Figure 2 Crystal structure of the Shq1-specific domain (SSD). (**A**) Domain organization of Shq1. The CS domain and SSD are labelled. (**B**) Ribbon representation of the SSD structure in cross-eye stereo-view. The secondary structural elements are indicated. The central helix α8 is green, the outer helical circle is blue, the inner helical circle is orange and the C-terminal structural elements are red. The N- and C-termini are marked by 'N' and 'C'. Dots are used to represent the disordered loops. (**C**) Topology diagram of the SSD structure. (**D**) Sequence alignment of the SSD. Ninety Shq1 proteins were aligned, but only the *S. cerevisiae* (Sc) and *H. sapiens* (Hs) sequences are shown. Residues that are conserved in at least 95, 80 and 60% of the 90 Shq1 proteins are shaded by black, grey and light grey, respectively. The secondary structure elements observed in the free SSD structure are indicated above the alignment. Residues with solvent accessible surface buried by at least 30 and 10 $Å^2$ owing to Cbf5 association are marked by red solid and open circles, respectively.

Shq1 or the CS domain cannot support growth, consistent with the previous results (Singh *et al*, 2009). Moreover, removal of the entire SSD was lethal, indicating that the SSD is also essential for yeast viability. Among the point mutations that impair the Shq1–Cbf5 interaction in ITC experiments, R383E reduced cell growth at 30°C, but W326D showed nearly normal growth at 30°C. We further examined how temperature influences the growth phenotype of these mutations. Interestingly, the growth of R383E and W326D mutants was severely inhibited at 37°C, but were nearly normal at 20°C (Figure 5B). The temperature-dependent phenotype was not due to the reduced protein level of Shq1 mutants at high temperatures (Figure 5C). These results indicate that disruption of the Shq1–Cbf5 interaction is detrimental, particularly at high growth temperatures.

DC mutations and Shq1 binding

A majority of X-linked DC mutations are clustered on the N- and C-terminal regions of Cbf5, part of which constitute the PUA domain (Marrone *et al*, 2005). Compared with their archaeal counterparts, eukaryotic Cbf5 proteins contain the N-terminal extension (NTE) and CTE (Figure 6A). These two extensions are highly conserved in eukaryotes and harbour many DC mutations. The crystal structure of the CNG complex showed that the NTE folds into a new structural layer covering on the PUA domain, whereas the CTE is disordered (Li *et al*, 2011). The Shq1 complex structure now reveals a novel function for the CTE in mediating Shq1 association.

The Shq1 complex structure shows that yeast residues Pro354, Ala356, Leu368, Gly372, Thr378 and Pro379 in the CTE and K284 in the PUA domain, which correspond to DC mutation sites, are located at the Shq1-binding interface (Figure 6B). DC mutations at these sites may affect Shq1 binding and hence H/ACA RNP and telomerase biogenesis. The truncation mutant Cbf5Δ354, which lacks a large portion of the CTE, could be regarded as an extreme model for DC mutations in the CTE. As shown above (Figure 1B), Cbf5Δ354 shows a 4-fold weaker Shq1-binding affinity than Cbf5 with a complete CTE. A single site DC mutation in this region should display a milder effect on Shq1 binding than Cbf5Δ354.

However, most of DC mutation sites in the PUA domain are not located at the binding surface for Shq1 or RNA (Figure 6B). Nevertheless, the effect on Shq1 binding of mutations at these sites may be indirect. To experimentally test this notion, we used the yeast Cbf5 protein to assess whether nine DC mutations in the PUA domain affect Shq1 binding (Figure 6C). Seven sites were substituted with the exact amino acid found in DC patients (I8T, K9E, E11K, T19M, R35T, S36A and L291F), while the other two sites were replaced with amino acid with more different physicochem-

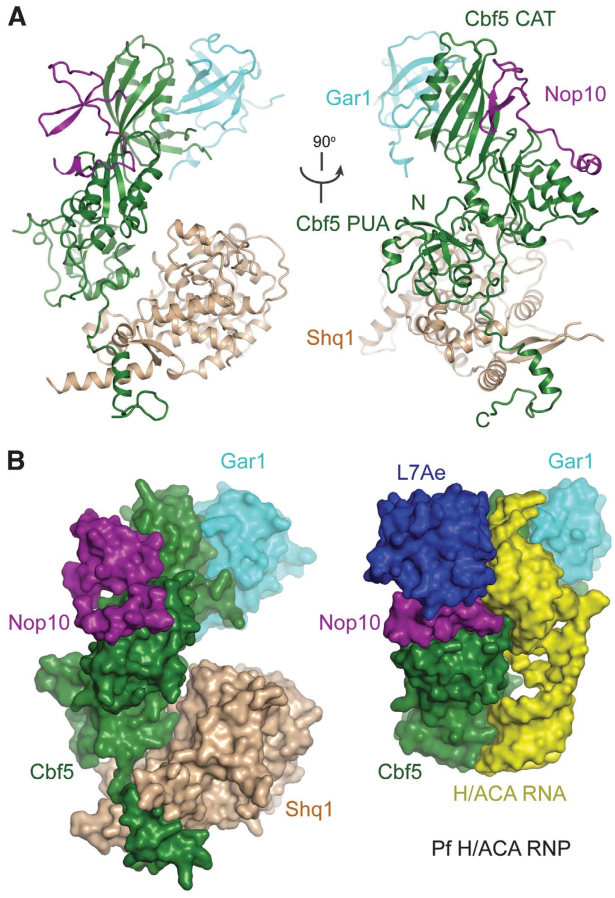

Figure 3 Structure of the SSD–Cbf5–Nop10–Gar1 complex. (**A**) Ribbon representation of the structure, shown in two orthogonal views. Cbf5 is green, Nop10 is magenta, Gar1 is cyan and Shq1 is wheat. (**B**) Surface view of the Shq1 complex and the *P. furiosus* H/ACA RNP structure. The two structures are aligned by the catalytic domain of Cbf5. The RNA is yellow.

ical properties (H38A and V323S). Moreover, we made a deletion mutant ΔN15 that removed the N-terminal 15 residues including 7 DC mutation sites.

These Cbf5 DC mutants were co-purified with Nop10, or with Nop10 and Gar1 for a better solubility, and assessed for Shq1 binding using the Ni-bead pull-down assay (Figure 6C). To assess the sensitivity of the pull-down assay, we included Cbf5Δ354 as a negative control, whose Shq1-binding affinity has been quantified with ITC. The results show that Cbf5Δ354 retained much less Shq1 compared with Cbf5 with an intact CTE (lane 4), indicating that the pull-down assay is sensitive enough to detect a 4-fold reduction in binding affinity. All of the DC mutants and ΔN15 retained stoichiometric amounts of Shq1 similarly to wild-type Cbf5. These results indicate that these DC mutations in the PUA domain of yeast Cbf5 do not appreciably affect Shq1 binding.

Discussion

Shq1 is an H/ACA RNP assembly chaperone

As an assembly factor, Shq1 is present neither in mature H/ACA RNPs nor in H/ACA pre-RNPs assembled at H/ACA RNA genes. The mechanism by which Shq1 is excluded from H/ACA RNPs has been ascribed to the presence of other Cbf5-associated proteins (Grozdanov *et al*, 2009b). We have shown, by biochemical and structural analysis, that the H/ACA RNP proteins Nop10, Nhp2 and Gar1 and the assembly factor Naf1 do not interfere with the Shq1–Cbf5 interaction *in vitro*. Rather, H/ACA RNAs could dislodge Shq1 from Cbf5, as they share an overlapping binding site on the PUA domain. Our data also suggest that Shq1 transfers the pre-assembled Cbf5 protein complexes to nascent H/ACA RNAs.

The discrepancy between our results and the previous studies (Grozdanov *et al*, 2009b) may be because the yeast proteins we used have species-specific differences compared with their rat counterparts, which were used previously. However, this possibility seems unlikely as the sequences of H/ACA RNP proteins are highly conserved in eukaryotes. Alternatively, the previous studies were performed in rabbit reticulocyte lysates or in mammalian cells and the endogenous H/ACA RNAs present in these systems might interfere with the Shq1–Cbf5 interaction.

The PUA domain of Cbf5 is a major RNA binding domain recognizing the lower stem and ACA motif of H/ACA RNAs (Li and Ye, 2006). The Shq1 blockage of the PUA domain would prevent illicit binding of non-specific RNAs before Cbf5 assembles with H/ACA RNAs. In this regard, Shq1 fulfils the role as an assembly chaperone. Assembly chaperones associate with assembly intermediates of many macromolecular complexes to pre-organize individual subunits and prevent pre-mature interactions or dead-end products (Chari and Fischer, 2010). Recently, two assembly chaperones involved in snRNP biogenesis were structurally illustrated. The Gemin2 protein of the SMN complex was shown to wrap around a Sm pentamer complex and block RNA binding (Zhang *et al*, 2011). Aar2, a U5 snRNP assembly factor, was shown to bind Prp8 and prevent the association of Brr2, which is present in mature U5 snRNP (Weber *et al*, 2011).

Assembly chaperones also trap the complexes in intermediate states that need to be disassembled before the next step of assembly (Chari and Fischer, 2010). The competitive binding of H/ACA RNA might provide a mechanism for Shq1 displacement from H/ACA RNPs. We analysed the Cbf5 binding competition between H/ACA RNA and Shq1, but obtained different results for two tested RNAs snR5 and snR36Δins, which is a derivative of snR36 lacking an insertion in the 5′ hairpin (unpublished observation). SnR36Δins, but not snR5, can displace Shq1 from the Cbf5–Nop10–Gar1–Nhp2 complex. In the case of snR36Δins, Shq1 displacement was reduced in the absence of Nhp2, suggesting that Nhp2, which presumably binds the upper stem of H/ACA RNA, enhances the competition ability of the H/ACA RNA. The two RNAs appear to possess different binding affinities to Cbf5 complexes. Moreover, direct RNA competition may not be the sole driving force for the *in-vivo* displacement of Shq1 and other factors likely exist to help disassemble the Shq1 complex.

We show that the phenotype of two Shq1 point mutations that impair Cbf5 interaction strongly depends on the growth temperature. The suppression of yeast growth was pronounced at 37°C, but largely disappeared at 20°C (Figure 5). The Shq1–Cbf5 interaction may be stabilized at low temperature, partially compensating the disruptive effect of the mutations. Alternatively, Shq1 may protect Cbf5 from misfolding and/or aggregation, which is expected to be more

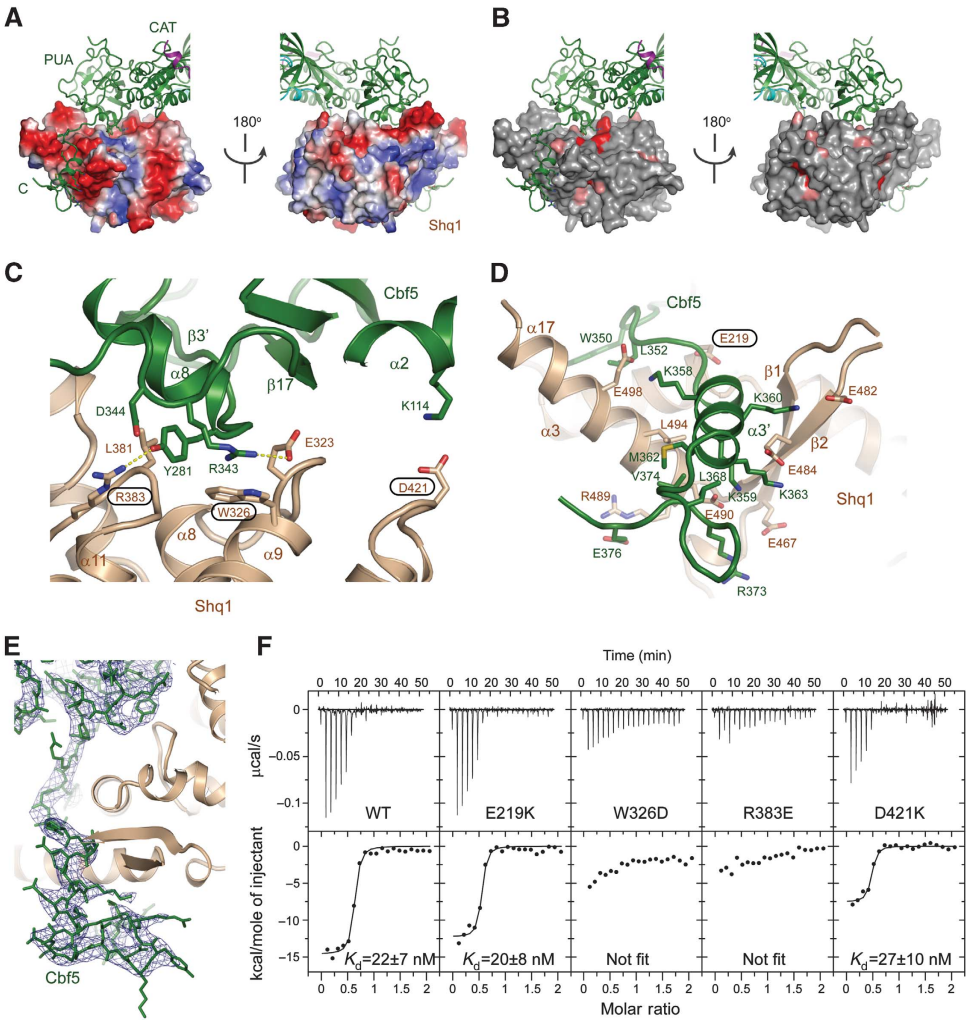

Figure 4 Interaction between Shq1 and Cbf5. (**A**) Electrostatic potential surface of the Shq1 domain in two opposite orientations. The surface is coloured from red to blue for negatively to positively charged regions. (**B**) Conserved surface of the Shq1 domain in two opposite orientations. Residues with at least 95 and 80% conservation as defined in Figure 2D are coloured in red and pink, respectively. (**C**) Interactions between Shq1 and the PUA domain of Cbf5. Interacting residues are shown as sticks with the oxygen atoms in red, the nitrogen atoms in blue and the carbon atoms in green for Cbf5 and in wheat for Shq1. Dashed lines denote hydrogen bonds. Residues analysed by mutagenesis are boxed. (**D**) Interactions between Shq1 and the C-terminal extension of Cbf5. (**E**) The $2f_o-f_c$ electron density map counted at the 1σ level is shown for Cbf5. (**F**) ITC profiles of Shq1 and mutants with the Cbf5–Nop10 complex. The K_d values are indicated.

severe at high temperature. Interestingly, Shq1 displays a general chaperone activity *in vitro* that suppresses thermally induced aggregation of citrate synthase in a stoichiometric manner (Godin *et al*, 2009). The CS domain of Shq1 is structurally similar to Hsp20 proteins, which usually bind to the Hsp90 chaperone and regulate its activity. However, no interaction has been detected between Hsp90 and the CS domain (Godin *et al*, 2009; Singh *et al*, 2009). Hsp90 has been shown to be involved in H/ACA RNP biogenesis (Boulon *et al*, 2008; Zhao *et al*, 2008). The exact role of Shq1 in H/ACA RNP biogenesis and its connection to other chaperones remain to be sorted out.

DC mutations and Shq1 interaction

Most DC mutants in dyskerin are clustered on the PUA domain and the eukaryote-specific NTE and CTE. The previous archaeal Cbf5 structure cannot be used to model many DC mutations in the extensions. The structures of yeast Cbf5 reveal the role of the NTE in building an enlarged eukaryotic PUA domain and the CTE in binding Shq1, providing a new framework for investigation of the effect of DC mutations.

We introduced a few DC mutations into the PUA domain of yeast Cbf5 but did not detect an effect on the Cbf5–Shq1 interaction. The mutational effects on Shq1 binding might be too subtle to be detected by our *in-vitro* method, but may remain significant *in vivo*. It is also possible that DC mutations affect the protein stability *in vivo* and the PUA domain is involved in additional unknown functions besides binding H/ACA RNA and Shq1. A previous study showed that several mutations at DC sites in the PUA domain modulate the interaction between rat MBP-NAP57 and human SHQ1 (Grozdanov *et al*, 2009b). It should be noted that we measured the binding between yeast Shq1 and Cbf5 in 400 mM salt, while in such a concentration of salt the interaction between rat MBP-NAP57 and human SHQ1 was largely disrupted (Grozdanov *et al*, 2009a). The different properties of the proteins used may lead to some difference in our and the previous studies.

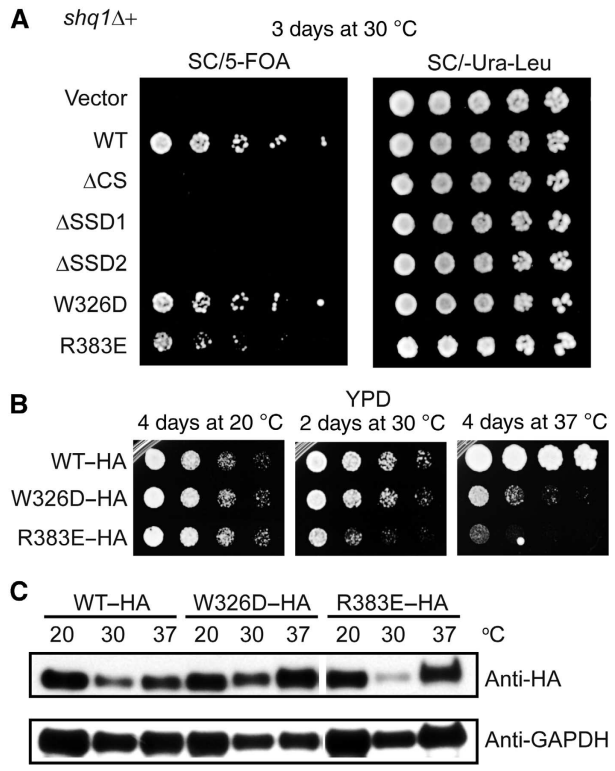

Figure 5 The Shq1–Cbf5 interaction is important for yeast growth. (**A**) Yeast growth assay of *SHQ1* mutants. An *shq1*Δ haploid strain complemented by wild-type *SHQ1* on a *URA3* plasmid was transformed with an empty *LEU2* plasmid (vector), or *LEU2* plasmids encoding wild-type *SHQ1* gene (WT), *SHQ1* residues 100–507 (ΔCS), residues 1–100 (ΔSSD1), residues 1–146 (ΔSSD2) or *SHQ1* with point mutations W326D or R383E. The transformants were grown on synthetic complete (SC) medium with 5FOA to counter-select for the wild-type *SHQ1 URA3* plasmid and on Ura- and Leu-deficient SC media (SC-Ura-Leu) at 30°C. (**B**) Strains with HA-tagged *SHQ1* or its W326D and R383E mutants were generated as in (**A**) and recovered from 5FOA plates and further grown in YPD plates at 20, 30 and 37°C for indicated time. (**C**) The HA-tagged *SHQ1* strains were grown at indicated temperatures and analysed for protein level by immunoblotting with anti-HA and anti-GAPDH (loading control) antibodies.

Materials and methods

DNA constructs and protein purification

The sequence encoding full-length Shq1 (residues 1–507) was amplified by PCR from yeast genomic DNA and cloned into the first multiple cloning site of a modified pETDuet-1 plasmid (Novagen). The gene encoding the SSD (residues 147–507) was amplified from the full-length *Shq1* gene and cloned similarly. The Shq1 proteins were expressed as fusion proteins with an N-terminal 6-histidine tag, thioredoxin (Trx) and a linker with a PreScission cleavage site. Point mutants were generated by the QuikChange method (Stratagene) and confirmed by DNA sequencing.

Shq1 and the SSD were expressed in the *E. coli* Rosetta 2 (DE3) strain (Novagen). The expression bacteria were cultured in LB medium containing 50 μg ml^{-1} ampicillin and 34 μg ml^{-1} chloramphenicol at 37°C to an OD_{600} of 1.0; 1 mM IPTG was then added for protein induction overnight at 18°C. The cells harvested from 2 L of culture were resuspended in 100 ml buffer A (50 mM sodium phosphate (pH 7.6) and 300 mM NaCl), lysed with a high-pressure cell disruptor and centrifuged to remove the cell debris. The supernatant was applied to a 5-ml HisTrap column (GE Healthcare) equilibrated in buffer A. The column was washed with 50 ml buffer A and with 50 ml buffer A supplemented with 20 mM imidazole. The target protein was eluted with 500 mM imidazole in buffer A and supplemented with 2 mM dithiothreitol. To remove the N-terminal His-Trx tag, the protein was incubated with recombinant PreScission protease overnight at 4°C. The digested protein was diluted by adding two volumes of buffer B (25 mM HEPES-Na (pH 7.6)), loaded onto a 5-ml HisTrap Q column (GE Healthcare) pre-equilibrated with buffer B and eluted with a 0–1 M NaCl gradient. Fractions containing the target protein were pooled, concentrated to 2 ml and further purified with a Superdex S200 16/60 column equilibrated in 10 mM HEPES-K (pH 7.6) and 100 mM NaCl. The protein was concentrated to 40 mg ml^{-1}, divided into aliquots and stored at −80°C.

The SSD was labelled with selenomethionine in M9 medium by blocking methionine biosynthesis (Van Duyne *et al*, 1993). Se-labelled protein was purified by the same procedure as unlabelled protein except that the lysis buffer was supplemented with 10 mM β-mercaptoethanol.

For purification of the Shq1–Cbf5–Nop10–Gar1 complex, the CNG complex was purified by HisTrap and heparin chromatography as described (Li *et al*, 2011). The His tag of Gar1 was not cleaved. The SSD was purified after Q chromatography as described above. The CNG complex was mixed with the SSD in a 1:3 molar ratio and incubated on ice overnight. The quaternary complex was separated from excessive free SSD on a Superdex S200 16/60 column.

Crystallization

The crystal of the SSD alone was grown using the hanging-drop vapour-diffusion method by mixing 1 μl of 40 mg ml^{-1} Shq1 (147–507) in 10 mM HEPES-K (pH 7.5) and 100 mM NaCl and 1 μl of reservoir solution containing 0.1 M tri-sodium citrate (pH 6.0), 0.2 M ammonium acetate, 20% w/v PEG 4000 and 9% isopropanol. Rod-shaped crystals appeared in 2 days for the native protein at 20°C and in 2 weeks for the Se-labelled protein at 4°C. The crystals were cryoprotected in 20% glycerol made in the reservoir solution and flash frozen in liquid nitrogen. The crystal of the Shq1–Cbf5–Nop10–Gar1 complex (20 mg ml^{-1} in 50 mM Tris–Cl (pH 8.0), 500 mM NaCl, 10% glycerol and 1 mM DTT) was grown in 0.1 M Bis-Tris (pH 6.5) and 20% w/v polyethylene glycol monomethyl ether 5000 at 20°C and cryoprotected in 16% glycerol in the reservoir solution.

Data collection and structure determination

Diffraction data were collected at the Shanghai Synchrotron Radiation Facility beamline BL17U1 and processed by HKL2000 (Otwinowski and Minor, 1997). The crystal of the SSD belongs to space group P1 and contains one molecule per asymmetric unit. The SSD structure was solved by the isomorphous replacement with anomalous scattering method based on a 1.6-Å native data set and a 2.1-Å selenium derivative data set collected at a wavelength of 0.9793 Å. Heavy atoms were identified with SHELXD (Sheldrick, 2008). The phases were calculated with SHARP and solvent modified (Vonrhein *et al*, 2007). The resultant experimental map was of excellent quality and allowed automatically building most of the model in PHENIX (Adams *et al*, 2010). The model was further manually adjusted in Coot (Emsley and Cowtan, 2004) and refined in Refmac (CCP4, 1994; Murshudov *et al*, 1999). The current model of the SSD includes residues 164–345, 363–471 and 483–503 and 330 water molecules. The RAMPAGE analysis showed that 98.4% of the residues are in favourable region, 1.3% are in allowed region and 0.3% are in outlier region (Lovell *et al*, 2003). A single residue (Tyr379) falls in the outlier region of Pre-proline and displays well-defined election density.

The crystal of the Shq1–Cbf5–Nop10–Gar1 complex belongs to space group $P4_32_12$ and contains one tetramer per asymmetric unit. The structure was solved by molecular replacement in Phaser (McCoy *et al*, 2007), using the CNG complex structure and the SSD structure as search models (Li *et al*, 2011). The structure was refined in Refmac to 3.06 Å and includes Shq1 (residues 164–347, 361–476 and 480–506), Cbf5 (18–156 and 161–378), Nop10 (1–48) and Gar1 (33–124). The structural figures were created in PyMOL (DeLano, 2002).

Pull-down assay

The Cbf5 complexes of various compositions were co-expressed and co-purified, and Naf1 (107–280) was purified alone, as described previously (Li *et al*, 2011). Naf1 (240 pmol) was assembled with 120 pmol Cbf5 complexes. Among these proteins, only Cbf5 had a His tag, while Nop10 lacked a His tag, and the His tags of Gar1, Naf1 and Nhp2 were cleaved during purification. Each of the Cbf5 complexes (120 pmol) was mixed with a 4-fold molar excess of Shq1 (480 pmol) and 20 μl Ni beads in 500 μl binding buffer containing

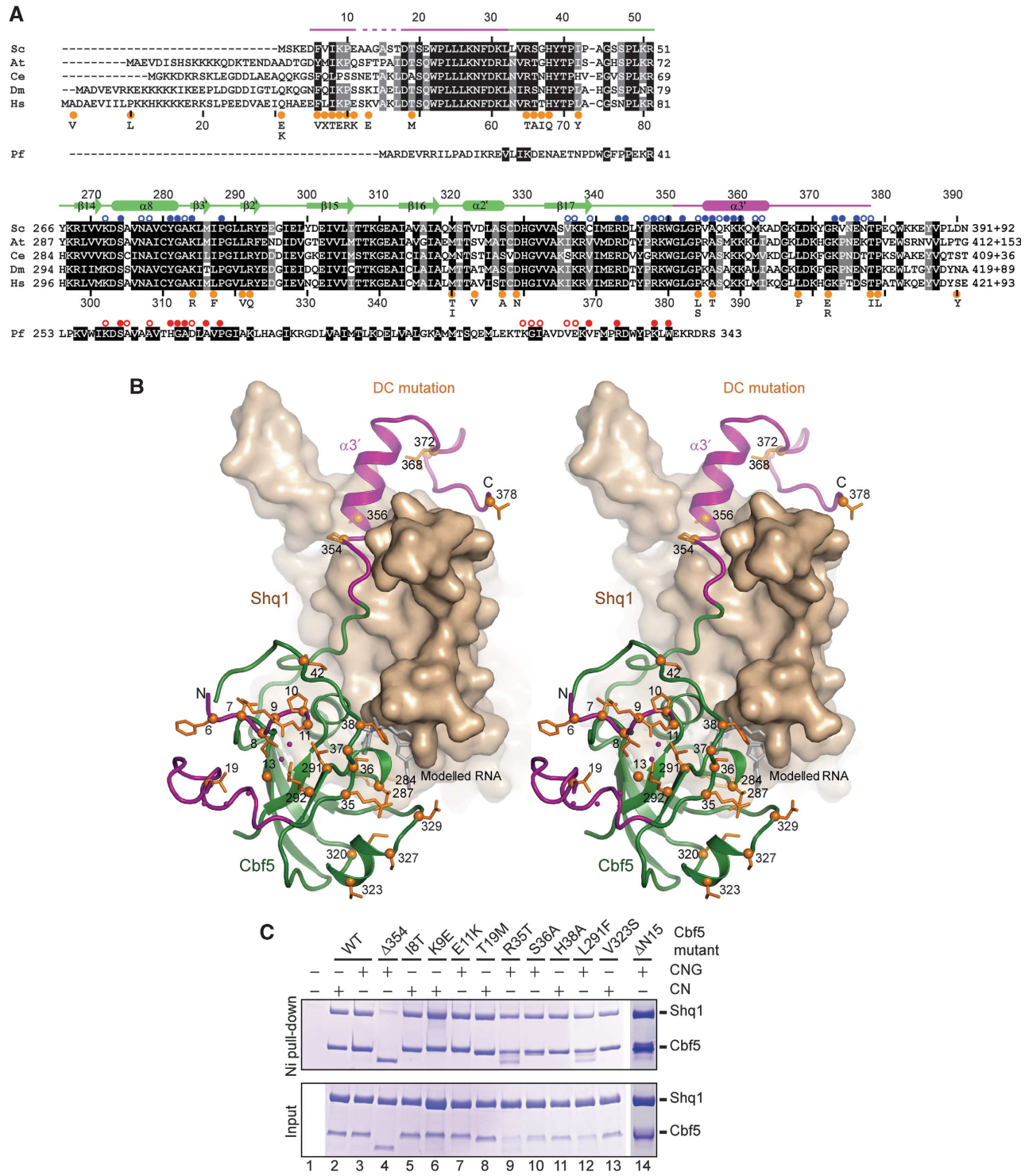

Figure 6 DC mutations and Shq1 interaction. (**A**) Sequence alignment around the Cbf5 PUA domain. The aligned Cbf5 sequences are from *S. cerevisiae* (Sc), *A. thaliana* (At), *C. elegans* (Ce), *D. melanogaster* (Dm), *H. sapiens* (Hs) and *P. furiosus* (Pf). Residues that are conserved in 100 and 80% of the five eukaryotic sequences are shaded by black and grey, respectively. Pf residues that match the conserved eukaryotic residues are shaded. The observed secondary structure elements are indicated with the NTE and CTE coloured in magenta. DC mutations are labelled and denoted by orange dots below the human dyskerin sequence (X for deletion). Yeast Cbf5 residues with solvent accessible surface buried by at least 30 and 10 $Å^2$ due to Shq1 association are marked by blue solid and open circles, respectively. Pf Cbf5 residues buried by H/ACA RNA are marked by red circles in a similar manner. (**B**) DC mutations mapped to the PUA domain and the C-terminal extension of the yeast Cbf5–Shq1 structure (cross-eye stereo-view). The DC mutation sites are indicated by orange spheres for Cα atoms and sticks, and labelled with the residue number of yeast Cbf5. For the purposes of illustration, residues 5–11 not observed in the Shq1 complex structure are modelled as in the CNG structure. (**C**) Interaction between Shq1 and yeast Cbf5 with DC mutations. The Cbf5 DC mutant complexes co-purified with either Nop10 (CN) or Nop10 and Gar1 (CNG) were incubated with a 4-fold excess of Shq1 and pulled down with Ni beads. The input panel contained 5% of the material used for the pull-down. Only the portion of gel containing Cbf5 and Shq1 is shown. Δ354 and ΔN15 contains Cbf5 residues 1–354 and 16–394, respectively.

50 mM sodium phosphate (pH 7.6) and 0.4 M KCl. For Cbf5 with DC mutations (Figure 6C), the binding buffer contained 50 mM sodium phosphate (pH 7.6), 0.4 M NaCl and 5% glycerol. The binding reactions were nutated gently for 2 h at 4°C. The Ni beads were collected by centrifugation and washed once with 1 ml 40 mM imidazole in binding buffer. Bound proteins were eluted with 20 μl 500 mM imidazole in binding buffer. For analysis by SDS–PAGE, 10 μl of the eluate was mixed with 10 μl 2 × SDS-loading buffer.

ITC

ITC titrations were performed at 25°C using an ITC-200 microcalorimeter (MicroCal Inc.). All proteins were exchanged into 50 mM sodium phosphate (pH 7.6) and 0.4 M KCl buffer using a Superdex 200 16/60 column and were supplemented with 0.01% NP-40 prior to the ITC experiment. Full-length Shq1 or its mutants (50 or 100 μM) were titrated into the Cbf5 complex (5 or 10 μM) in a 200-μl sample cell by a 40-μl microsyringe as follows: 0.4 μl for the first injection and 2 μl for the next 19 injections with an interval of 150 s. The integrated heat data were analysed using the one-set-of-sites model in MicroCal Origin according to the manufacturer's instructions. The first data point was not used in the analysis. The binding parameters ΔH (reaction enthalpy change in cal mol^{-1}), K (binding constant in M^{-1}) and n (number of bound Shq1 molecules per Cbf5 complex) were floating during the fit. The binding free energy ΔG and reaction entropy ΔS were calculated using the relationships $\Delta G = -RT\ln K$ ($R = 1.9872$ cal mol^{-1} K^{-1}, $T = 298$ K) and $\Delta G = \Delta H - T\Delta S$. The dissociation constant K_d was calculated as $1/K$. The derived thermodynamic parameters are listed in Supplementary Table S1.

Yeast plasmids and strains

The wild-type *SHQ1* gene, including 225 bp sequences upstream of the start codon and 300 bp sequences downstream of the stop codon, was amplified by PCR from yeast genomic DNA and cloned into plasmid pRS416 with a *URA3* marker to yield pRS416-SHQ1 and into pRS415 with a *LEU2* marker to yield pRS415-SHQ1 (Brachmann *et al*, 1998). Single point mutations (W326D or R383E) and domain deletions (ΔCS, ΔSSD1 or ΔSSD2) were introduced into pRS415-SHQ1 with the QuikChange method. For domain deletions, the whole DNA region encoding the domain of interest was removed from the plasmid. The HA tag was inserted to the C-terminus of *SHQ1* gene by QuikChange.

The heterozygous *Shq1* deletion diploid strain BY4743 *SHQ1/shq1Δ::kanMX4* (*MAT* a/α *his3Δ1/his3Δ1*, *leu2Δ0/leu2Δ0, lys2Δ0/LYS2, MET15/met15Δ0, ura3Δ0/ura3Δ0, SHQ1/shq1Δ::kanMX4*) was purchased from Euroscarf and was maintained in YPD medium complemented with 2% glucose and 0.003% adenine. The strain was transformed with pRS416-SHQ1, and the transformants were sporulated and dissected. The spores were germinated into colonies on YPD plates. Each clone was duplicated on Ura-deficient synthetic complete (SC) medium supplemented with antibiotic G418 to select the *shq1Δ* haploid clone with *URA3* pRS416-SHQ1. The surviving clones were further examined for the absence of genomic *SHQ1* copy by PCR. Growth and handling of yeast were performed by standard techniques. Yeast cells were transformed using the PEG/lithium acetate method. The constructed plasmids and yeast strains and used oligonucleotides are listed in Supplementary Tables S2–S4, respectively.

Growth assay

The *shq1Δ* haploid strain with *URA3* pRS416-SHQ1 was transformed with *LEU2* pRS415-SHQ1 containing either wild-type or mutant *SHQ1*. The transformants were grown in 2 ml Ura- and Leu-deficient SC medium for 16 h at 30°C and adjusted to an OD_{600} of 0.5. The culture was serially diluted by 4-fold in 96-well plates and spotted with a 48-pin replicator onto SC plates with 5-fluoroorotic acid (5FOA) to counter select the *URA3* pRS416-SHQ1 plasmid or without 5FOA. The plates were incubated for 3 days at 30°C. The stains with HA-tagged *SHQ1* and its W326D and R383E mutant were recovered from SC/5FOA plates, cultured in liquid YPD medium and spotted onto YPD plates for 2–4 days at 20, 30 and 37°C.

Immunoblot

For analysis of protein stability, the SHQ1–HA strains were grown in YPD medium at 20, 30 and 37°C to about 1.5 OD at 600 nm. Cells of 4 OD were collected and resuspended in 150 μl lysis buffer containing 40 mM Tris–Cl (pH 6.8), 5% w/v SDS, 8 M urea, 0.1 mM EDTA, 0.4 mg ml^{-1} bromophenol blue and 1 mM PMSF. The sample was mixed with 80 μl glass beads (0.5 mm in diameter, Biospec) and lysed with four times of 30-s vortex. The clarified supernatant (10 μl) was resolved in a 4–12% gradient SDS–PAGE gel and immunoblotted with anti-HA (Santa Cruz Biotechnology) and anti-GAPDH antibodies.

Accession codes

The atomic coordinates and structure factors have been deposited in the Protein Data Bank under accession number 3UAH for the Shq1-specific domain and 3UAI for the Shq1–Cbf5–Nop10–Gar1 complex structure.

Supplementary data

Supplementary data are available at *The EMBO Journal* Online (http://www.embojournal.org).

Acknowledgements

We thank the staff at the Shanghai Synchrotron Radiation Facility beamline BL17U1 for assistance in data collection. We thank Drs Jinqiu Zhou and Li-Lin Du for help in yeast experiments. This research was supported by the Chinese Ministry of Science and Technology through 973 (2010CB835402) and 863 (2008AA022310) projects and the Beijing Municipal Government.

Author contributions: SL, JD and KY designed experiments, analysed data and wrote the manuscript. SL, JD, DL, SM and KY performed the experiments.

Conflict of interest

The authors declare that they have no conflict of interest.

References

Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung LW, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner R, Read RJ, Richardson DC, Richardson JS, Terwilliger TC, Zwart PH (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66:** 213–221

Baker DL, Youssef OA, Chastkofsky MI, Dy DA, Terns RM, Terns MP (2005) RNA-guided RNA modification: functional organization of the archaeal H/ACA RNP. *Genes Dev* **19:** 1238–1248

Ballarino M, Morlando M, Pagano F, Fatica A, Bozzoni I (2005) The cotranscriptional assembly of snoRNPs controls the biosynthesis of H/ACA snoRNAs in Saccharomyces cerevisiae. *Mol Cell Biol* **25:** 5396–5403

Batista LF, Pech MF, Zhong FL, Nguyen HN, Xie KT, Zaug AJ, Crary SM, Choi J, Sebastiano V, Cherry A, Giri N, Wernig M, Alter BP, Cech TR, Savage SA, Reijo Pera RA, Artandi SE (2011) Telomere shortening and loss of self-renewal in dyskeratosis congenita induced pluripotent stem cells. *Nature* **474:** 399–402

Bessler M, Wilson DB, Mason PJ (2010) Dyskeratosis congenita. *FEBS Lett* **584:** 3831–3838

Boulon S, Marmier-Gourrier N, Pradet-Balade B, Wurth L, Verheggen C, Jady BE, Rothe B, Pescia C, Robert MC, Kiss T, Bardoni B, Krol A, Branlant C, Allmang C, Bertrand E, Charpentier B (2008) The Hsp90 chaperone controls the biogenesis of L7Ae RNPs through conserved machinery. *J Cell Biol* **180:** 579–595

Brachmann CB, Davies A, Cost GJ, Caputo E, Li J, Hieter P, Boeke JD (1998) Designer deletion strains derived from Saccharomyces cerevisiae S288C: a useful set of strains and plasmids for PCR-mediated gene disruption and other applications. *Yeast* **14:** 115–132

CCP4 (1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr D Biol Crystallogr* **50:** 760–763

Chari A, Fischer U (2010) Cellular strategies for the assembly of molecular machines. *Trends Biochem Sci* **35:** 676–683

Charpentier B, Muller S, Branlant C (2005) Reconstitution of archaeal H/ACA small ribonucleoprotein complexes active in pseudouridylation. *Nucleic Acids Res* **33:** 3133–3144

Darzacq X, Jady BE, Verheggen C, Kiss AM, Bertrand E, Kiss T (2002) Cajal body-specific small nuclear RNAs: a novel class of 2′-O-methylation and pseudouridylation guide RNAs. *EMBO J* **21:** 2746–2756

Darzacq X, Kittur N, Roy S, Shav-Tal Y, Singer RH, Meier UT (2006) Stepwise RNP assembly at the site of H/ACA RNA transcription in human cells. *J Cell Biol* **173:** 207–218

DeLano WL (2002) *The PyMOL User's Manual*. Delano Scientific: San Carlos, CA, USA

Dez C, Noaillac-Depeyre J, Caizergues-Ferrer M, Henry Y (2002) Naf1p, an essential nucleoplasmic factor specifically required for accumulation of box H/ACA small nucleolar RNPs. *Mol Cell Biol* **22:** 7053–7065

Duan J, Li L, Lu J, Wang W, Ye K (2009) Structural mechanism of substrate RNA recruitment in H/ACA RNA-guided pseudouridine synthase. *Mol Cell* **34:** 427–439

Emsley P, Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60:** 2126–2132

Fatica A, Dlakic M, Tollervey D (2002) Naf1p is a box H/ACA snoRNP assembly factor. *RNA* **8:** 1502–1514

Girard JP, Bagni C, Caizergues-Ferrer M, Amalric F, Lapeyre B (1994) Identification of a segment of the small nucleolar ribonucleoprotein-associated protein GAR1 that is sufficient for nucleolar accumulation. *J Biol Chem* **269:** 18499–18506

Godin KS, Walbott H, Leulliot N, van Tilbeurgh H, Varani G (2009) The box H/ACA snoRNP assembly factor Shq1p is a chaperone protein homologous to Hsp90 cochaperones that binds to the Cbf5p enzyme. *J Mol Biol* **390:** 231–244

Grozdanov PN, Fernandez-Fuentes N, Fiser A, Meier UT (2009a) Pathogenic Nap57 mutations decrease ribonucleoprotein assembly in dyskeratosis congenita. *Hum Mol Genet* **18:** 4546–4551

Grozdanov PN, Roy S, Kittur N, Meier UT (2009b) SHQ1 is required prior to NAF1 for assembly of H/ACA small nucleolar and telomerase RNPs. *RNA* **15:** 1188–1197

Heiss NS, Knight SW, Vulliamy TJ, Klauck SM, Wiemann S, Mason PJ, Poustka A, Dokal I (1998) X-linked dyskeratosis congenita is caused by mutations in a highly conserved gene with putative nucleolar functions. *Nat Genet* **19:** 32–38

Hoareau-Aveilla C, Bonoli M, Caizergues-Ferrer M, Henry Y (2006) hNaf1 is required for accumulation of human box H/ACA snoRNPs, scaRNPs, and telomerase. *RNA* **12:** 832–840

Holm L, Sander C (1993) Protein structure comparison by alignment of distance matrices. *J Mol Biol* **233:** 123–138

Jady BE, Bertrand E, Kiss T (2004) Human telomerase RNA and box H/ACA scaRNAs share a common Cajal body-specific localization signal. *J Cell Biol* **164:** 647–652

Kiss T, Fayet-Lebaron E, Jady BE (2010) Box H/ACA small ribonucleoproteins. *Mol Cell* **37:** 597–606

Leulliot N, Godin KS, Hoareau-Aveilla C, Quevillon-Cheruel S, Varani G, Henry Y, Van Tilbeurgh H (2007) The box H/ACA RNP assembly factor Naf1p contains a domain homologous to Gar1p mediating its interaction with Cbf5p. *J Mol Biol* **371:** 1338–1353

Li L, Ye K (2006) Crystal structure of an H/ACA box ribonucleoprotein particle. *Nature* **443:** 302–307

Li S, Duan J, Li D, Yang B, Dong M, Ye K (2011) Reconstitution and structural analysis of the yeast box H/ACA RNA-guided pseudouridine synthase. *Genes Dev* **25:** 2409–2421

Lovell SC, Davis IW, Arendall III WB, de Bakker PI, Word JM, Prisant MG, Richardson JS, Richardson DC (2003) Structure validation by Calpha geometry: phi,psi and Cbeta deviation. *Proteins* **50:** 437–450

Marrone A, Walne A, Dokal I (2005) Dyskeratosis congenita: telomerase, telomeres and anticipation. *Curr Opin Genet Dev* **15:** 249–257

Matera AG, Terns RM, Terns MP (2007) Non-coding RNAs: lessons from the small nuclear and small nucleolar RNAs. *Nat Rev Mol Cell Biol* **8:** 209–220

McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ (2007) Phaser crystallographic software. *J Appl Crystallogr* **40:** 658–674

Meier UT (2005) The many facets of H/ACA ribonucleoproteins. *Chromosoma* **114:** 1–14

Mitchell JR, Cheng J, Collins K (1999) A box H/ACA small nucleolar RNA-like domain at the human telomerase RNA 3′ end. *Mol Cell Biol* **19:** 567–576

Mochizuki Y, He J, Kulkarni S, Bessler M, Mason PJ (2004) Mouse dyskerin mutations affect accumulation of telomerase RNA and small nucleolar RNA, telomerase activity, and ribosomal RNA processing. *Proc Natl Acad Sci USA* **101:** 10756–10761

Murshudov GN, Vagin AA, Lebedev A, Wilson KS, Dodson EJ (1999) Efficient anisotropic refinement of macromolecular structures using FFT. *Acta Crystallogr D Biol Crystallogr* **55:** 247–255

Otwinowski Z, Minor W (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol* **276:** 307–326

Pogacic V, Dragon F, Filipowicz W (2000) Human H/ACA small nucleolar RNPs and telomerase share evolutionarily conserved proteins NHP2 and NOP10. *Mol Cell Biol* **20:** 9028–9040

Richard P, Darzacq X, Bertrand E, Jady BE, Verheggen C, Kiss T (2003) A common sequence motif determines the Cajal body-specific localization of box H/ACA scaRNAs. *EMBO J* **22:** 4283–4293

Robart AR, Collins K (2010) Investigation of human telomerase holoenzyme assembly, activity, and processivity using disease-linked subunit variants. *J Biol Chem* **285:** 4375–4386

Sheldrick GM (2008) A short history of SHELX. *Acta Crystallogr A* **64:** 112–122

Singh M, Gonzales FA, Cascio D, Heckmann N, Chanfreau G, Feigon J (2009) Structure and functional studies of the CS domain of the essential H/ACA ribonucleoparticle assembly protein SHQ1. *J Biol Chem* **284:** 1906–1916

Trahan C, Dragon F (2009) Dyskeratosis congenita mutations in the H/ACA domain of human telomerase RNA affect its assembly into a pre-RNP. *RNA* **15:** 235–243

Trahan C, Martel C, Dragon F (2010) Effects of dyskeratosis congenita mutations in dyskerin, NHP2 and NOP10 on assembly of H/ACA pre-RNPs. *Hum Mol Genet* **19:** 825–836

Van Duyne GD, Standaert RF, Karplus PA, Schreiber SL, Clardy J (1993) Atomic structures of the human immunophilin FKBP-12 complexes with FK506 and rapamycin. *J Mol Biol* **229:** 105–124

Venteicher AS, Meng Z, Mason PJ, Veenstra TD, Artandi SE (2008) Identification of ATPases pontin and reptin as telomerase components essential for holoenzyme assembly. *Cell* **132:** 945–957

Vonrhein C, Blanc E, Roversi P, Bricogne G (2007) Automated structure solution with autoSHARP. *Methods Mol Biol* **364:** 215–230

Walne AJ, Dokal I (2009) Advances in the understanding of dyskeratosis congenita. *Br J Haematol* **145:** 164–172

Weber G, Cristao VF, de LAF, Santos KF, Holton N, Rappsilber J, Beggs JD, Wahl MC (2011) Mechanism for Aar2p function as a U5 snRNP assembly factor. *Genes Dev* **25:** 1601–1612

Yang PK, Hoareau C, Froment C, Monsarrat B, Henry Y, Chanfreau G (2005) Cotranscriptional recruitment of the pseudouridylsynthetase Cbf5p and of the RNA binding protein Naf1p during H/ACA snoRNP assembly. *Mol Cell Biol* **25:** 3295–3304

Yang PK, Rotondo G, Porras T, Legrain P, Chanfreau G (2002) The Shq1p.Naf1p complex is required for box H/ACA small nucleolar ribonucleoprotein particle biogenesis. *J Biol Chem* **277:** 45235–45242

Zhang R, So BR, Li P, Yong J, Glisovic T, Wan L, Dreyfuss G (2011) Structure of a key intermediate of the SMN complex reveals Gemin2's crucial function in snRNP assembly. *Cell* **146:** 384–395

Zhao R, Kakihara Y, Gribun A, Huen J, Yang G, Khanna M, Costanzo M, Brost RL, Boone C, Hughes TR, Yip CM, Houry WA (2008) Molecular chaperone Hsp90 stabilizes Pih1/Nop17 to maintain R2TP complex activity that regulates snoRNA accumulation. *J Cell Biol* **180:** 563–578

Zhong F, Savage SA, Shkreli M, Giri N, Jessop L, Myers T, Chen R, Alter BP, Artandi SE (2011) Disruption of telomerase trafficking by TCAB1 mutation causes dyskeratosis congenita. *Genes Dev* **25:** 11–16

