

---

# Structural and functional analysis of Utp23, a yeast ribosome synthesis factor with degenerate PIN domain

---

JING LU,<sup>1,2</sup> MENGYI SUN,<sup>1,3</sup> and KEQIONG YE<sup>1,2,4</sup>

<sup>1</sup>National Institute of Biological Sciences, Beijing, Beijing 102206, China

<sup>2</sup>Graduate School of Peking Union Medical College and Chinese Academy of Medical Sciences, Beijing 100730, China

<sup>3</sup>College of Biological Sciences, China Agricultural University, Beijing 100193, China

## ABSTRACT

During synthesis of yeast ribosome, a large complex, called the 90S pre-ribosome or the small subunit processome, is assembled on the nascent precursor rRNA and mediates early processing of 18S rRNA. The Utp23 protein and snR30 H/ACA snoRNA are two conserved components of 90S pre-ribosomes. Utp23 contains a degenerate PIN nuclease domain followed by a long C-terminal tail and associates specifically with snR30. Here, we report the crystal structure of the Utp23 PIN domain at 2.5-Å resolution. The structure reveals a conserved core fold of PIN domain with degenerate active site residues, a unique CCHC Zn-finger motif, and two terminal extension elements. Functional sites of Utp23 have been examined with conservation analysis, mutagenesis, and *in vivo* and *in vitro* assays. Mutations in each of three cysteine ligands of zinc, although not the histidine ligand, were lethal or strongly inhibitory to yeast growth, indicating that the Zn-finger motif is required for Utp23 structure or function. The N-terminal helix extension harbors many highly conserved basic residues that mostly are critical for growth and *in vitro* RNA-binding activity of Utp23. Deletion of the C-terminal tail, which contains a short functionally important sequence motif, disrupted the interaction of Utp23 with snR30 and perturbed the pre-ribosomal association of Utp23. Our data establish a structural framework for dissecting Utp23 function in the assembly and dynamics of 90S pre-ribosomes.

**Keywords:** ribosome synthesis; crystal structure; PIN domain; 90S pre-ribosome

## INTRODUCTION

Eukaryotic ribosome synthesis is a complex and dynamic process that involves transcription, modification, and processing of precursor rRNA (pre-rRNA), coordinated assembly of ribosomal proteins, and export of pre-ribosomal particles from the nucleolus to the cytoplasm (for review, see Tschochner and Hurt 2003; Henras et al. 2008; Kressler et al. 2010; Panse and Johnson 2010; Karbstein 2011; Phipps et al. 2011). In the yeast *Saccharomyces cerevisiae*, ~200 *trans*-acting nonribosomal proteins and 75 small nucleolar RNAs (snoRNAs) have been found to participate in ribosome synthesis. This process starts in the nucleolus, where the small subunit 18S rRNA and the large subunit 5.8S and 25S rRNAs are transcribed by RNA polymerase I into a single polycistronic pre-rRNA. The nascent transcript undergoes extensive site-specific 2'-O-methylation and pseudouridylation with the modification site selected by C/D and H/ACA snoRNAs, respectively (Kiss et al. 2010; Watkins and Bohnsack 2011).

The primary transcript also contains 5'- and 3'-external transcribed spacers (5'-ETS and 3'-ETS) and internal tran-

scribed spacer 1 and 2 (ITS1 and ITS2) that are removed during a series of endo- and exonucleolytic cleavage. Early processing of 18S rRNA involves cleavage at sites A0 and A1 in the 5'-ETS and at site A2 in the ITS1, yielding the 20S pre-rRNA. The 20S pre-rRNA is exported to the cytoplasm and is processed at site D to produce mature 18S rRNA. Early processing of 18S rRNA occurs within the 90S pre-ribosomal particle, or the small subunit (SSU) processome, which are assembled cotranscriptionally on the 5' part of the primary pre-rRNA (Dragon et al. 2002; Grandi et al. 2002; Osheim et al. 2004). Genetic studies and biochemical purifications have found that roughly 70 nonribosomal proteins and four snoRNAs (U3, U14, snR10, snR30) are involved in early processing of 18S rRNA (Phipps et al. 2011). Many associate with 90S particles, but their binding strength or duration can be different.

Utp23 is a nucleolar protein that weakly associates with 90S pre-ribosomal particles and is required for early processing of 18S rRNA at sites A0, A1, and A2 (Bleichert et al. 2006).

---

<sup>4</sup>Corresponding author

E-mail yekeqiong@nibs.ac.cn

Article published online ahead of print. Article and publication date are at <http://www.rnajournal.org/cgi/doi/10.1261/rna.040808.113>.

© 2013 Lu et al. This article is distributed exclusively by the RNA Society for the first 12 months after the full-issue publication date (see <http://rnajournal.cshlp.org/site/misc/terms.xhtml>). After 12 months, it is available under a Creative Commons License (Attribution-NonCommercial 3.0 Unported), as described at <http://creativecommons.org/licenses/by-nc/3.0/>.

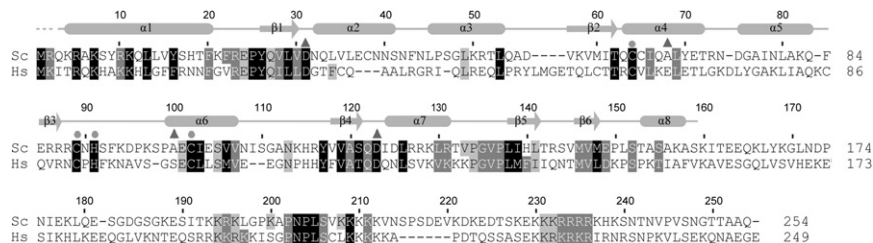
Utp23 contains a predicted PIN (Pit N terminus) domain. Typically, the PIN domain is an endoribonuclease of approximately 130 amino acids and is involved in various aspects of RNA processing and degradation (Clissold and Ponting 2000; Glavan et al. 2006; Arcus et al. 2011). The active site of the PIN domain consists of three or four acidic residues that bind  $Mg^{2+}$  or  $Mn^{2+}$  ions and catalyze the reaction. However, Utp23 appears to be an inactive nuclease because its equivalent active site residues are either mutated or dispensable for function (Bleichert et al. 2006). Synthesis of small ribosomal subunit also requires two other PIN domain proteins, Utp24 and Nob1. Utp24 has been proposed to be the endonuclease for A1 and/or A2 cleavage, and its active residues are essential for 18S rRNA processing at the two sites (Bleichert et al. 2006). Nob1 cleaves the 20S pre-rRNA at site D, generating the 3'-end of 18S rRNA (Fatica et al. 2003, 2004; Pertschy et al. 2009).

Utp23, along with another protein Kri1, was recently shown to associate specifically with the snR30 snoRNA (Hoareau-Aveilla et al. 2011). snR30 is a universally conserved H/ACA snoRNA that lacks a modification target, yet is required for pre-rRNA cleavage at A0, A1, and A2 sites (Morrissey and Tollervey 1993; Atzorn et al. 2004). The 3' hairpin of snR30 binds with two short separated sequences in the eukaryotic specific extension segment 6 of 18S rRNA (Fayet-Lebaron et al. 2009). Under normal conditions, only a small fraction of snR30 cosediments with large 90S pre-ribosomes in sucrose density gradients, suggesting that the association of snR30 with 90S pre-ribosomes is weak or transitive. Utp23 has been suggested to regulate snR30 release as depletion of Utp23 led to the accumulation of snR30 in large 90S pre-ribosomes (Hoareau-Aveilla et al. 2011). In this study, we have determined the crystal structure of the Utp23 PIN domain and identified functionally important sites of Utp23 with mutagenesis and a number of functional assays.

## RESULTS

### Sequence analysis and structural determination

To gain insight into the structural and functional organization of Utp23, we analyzed its sequence conservation. We searched homologs of *S. cerevisiae* Utp23 from the NCBI Reference Sequence (RefSeq) database with BLAST. Most sequenced eukaryotic genomes returned a single significant hit ( $E < 1 \times 10^{-10}$ ), which likely represents the ortholog of Utp23. Alignment of 134 sequences showed that Utp23 contains a conserved N-terminal region (residues 1–159), which corresponds to the predicted PIN domain, and a long C-terminal



**FIGURE 1.** Multiple sequence alignment of Utp23. A total of 134 Utp23 sequences were retrieved from RefSeq with BLAST and aligned with ClustalW and manual adjustment. Only sequences from *Saccharomyces cerevisiae* (Sc) and *Homo sapiens* (Hs) are displayed. Residues that are conserved in 95%, 80%, and 70% of all aligned sequences are shaded black, gray, and light gray, respectively. Similarity groups are defined as follows: D and E; K and R; S and T; F, Y, and W; and I, L, M, and V. The secondary structures observed in the crystal structure are indicated on the top. Triangles mark the equivalents to catalytic residues in PIN nucleases and circles denote Zn-coordinating residues.

region (residues 160–254) composed mainly of low complexity sequences (Fig. 1).

We expressed, purified, and crystallized full-length Utp23 protein and several C-terminally truncated fragments. A fragment that includes residue 1–159, covering only the PIN domain, yielded best crystals. We found from the sequence alignment that Utp23 contains four invariant residues—Cys64, Cys89, His91 and Cys102—in the PIN domain (Fig. 1). Since cysteine and histidine are frequent coordinating residues for metal ions such as Zn and Fe, we suspected that Utp23 may be a metal-binding protein. Indeed, characteristic radiation of Zn was detected in Utp23 crystals with X-ray fluorescence scanning at synchrotron. We therefore collected X-ray diffraction data at the peak absorption wavelength of Zn and determined the structure with single-wavelength anomalous dispersion (SAD) phasing. The structure was refined at 2.5-Å resolution to an  $R_{\text{free}}$ -value of 0.288 and an  $R_{\text{work}}$ -value of 0.228 (Table 1). The crystal belongs to space group P3<sub>2</sub>21 and contains two Utp23 molecules per asymmetric unit.

### Structure of the Utp23 PIN domain

In the crystal structure, each of two molecules inserts their C-terminal part (residues 144–159,  $\beta 6$ , and  $\alpha 8$ ) into the other molecule, forming a domain-swapped dimer (Fig. 2A,B). Other than that, the two molecules have little contact. Gel filtration chromatography showed that Utp23 1–159 is predominantly monomeric in solution (Supplemental Fig. S1), suggesting that the observed dimer in crystal is not the physiologically relevant conformation. Domain-swapped dimers have been observed for other crystallized proteins, for example, the FHA domain of Chfr mitotic checkpoint protein (Stavridi et al. 2002), yet their physiologic function is often unclear. High protein concentration used in crystallization could induce formation of a small fraction of thermodynamically equivalent domain-swapped dimer, which sometimes crystallizes readily.

**TABLE 1.** Data collection and refinement statistics

Data	Zn peak	Zn remote
Data collection		
Space group	P3 <sub>2</sub> 21	P3 <sub>2</sub> 21
Cell dimensions		
a, b, c (Å)	91.6, 91.6, 91.3	91.4, 91.4, 92.3
α, β, γ (°)	90, 90, 120	90, 90, 120
Wavelength (Å)	1.28174	0.97637
Resolution range (Å)	50–2.9 (2.95–2.9)	20–2.5 (2.54–2.5)
Unique reflections	10102 (521)	15536 (777)
Redundancy	12.2 (12.8)	8.2 (8.2)
<I>/<σ(I)>	45.6 (6.0)	30.5 (6.2)
Completeness (%)	99.8 (100)	99.6 (100)
R <sub>merge</sub>	0.107 (0.557)	0.103 (0.361)
Structure refinement		
Resolution range (Å)		20–2.51 (2.67–2.51)
Number of reflections		15380 (2508)
Number of atoms		2494
Mean B factors (Å <sup>2</sup> )		45.8
R <sub>work</sub>		0.228 (0.331)
R <sub>free</sub>		0.288 (0.409)
RMSD bond length (Å)		0.009
RMSD bond angles (°)		1.258

Values for the data in the highest resolution bin are shown in parentheses.

We describe below the monomeric structural model of Utp23 that is constructed with residues 3–143 of one molecule and residues 144–159 of the other molecule. The PIN domain folds into a three-layer  $\alpha/\beta/\alpha$  sandwich that consists a twisted six-stranded  $\beta$ -sheet at the center surrounded by eight  $\alpha$ -helices at both sides (Fig. 2C,D). The six strands are arranged in the order  $\beta_3$ - $\beta_2$ - $\beta_1$ - $\beta_4$ - $\beta_5$ - $\beta_6$ . All adjacent strands are parallel to each other, except that the C-terminal strand  $\beta_6$  is antiparallel to strand  $\beta_5$ .

We searched for structural homologs of Utp23 using the DALI server (Holm and Sander 1993) and found many PIN domain structures. These include the *Archaeoglobus fulgidus* protein AF0591 of unknown function ( $Z$ -score = 11.8) (Levin et al. 2004) and human SMG6 ( $Z$ -score = 8.3) (Glavan et al. 2006), which is involved in nonsense-mediated decay. The structure of the Utp23 PIN domain can be aligned onto other PIN domain structures with root mean square deviations (RMSDs) of 2–3 Å over approximately 100 Ca pairs. However, compared to canonical PIN domain structures, Utp23 additionally contains helix  $\alpha_1$  at the N terminus and strand  $\beta_6$  and helix  $\alpha_8$  at the C terminus.

As predicted, the structure shows that a Zn ion is coordinated to the S<sub>γ</sub> atoms of Cys64, Cys89, Cys102, and the N<sub>δ1</sub> atom of His91, forming a CCHC-type Zn finger (Fig. 2E). The Zn-finger motif brings together and stabilizes a local structure comprising helices  $\alpha_4$  and  $\alpha_6$  and the loop linking strand  $\beta_3$  and helix  $\alpha_6$ . Mutations of each of three Cys ligands to alanine were lethal or severely inhibited yeast growth, although mutation of the histidine ligand was neutral

(Fig. 3C). This indicates that the Zn-finger motif is important for maintaining Utp23 structure or function.

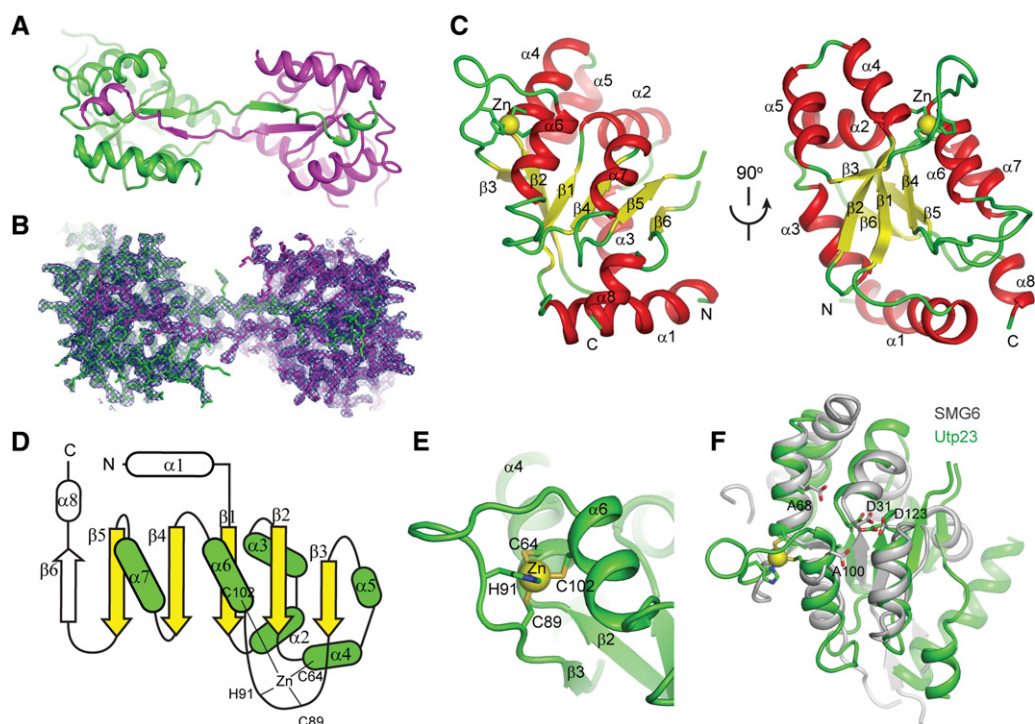
The classic PIN domain functions as endoribonuclease that cleaves single-stranded RNA in a Mg<sup>2+</sup> or Mn<sup>2+</sup>-dependent manner (Arcus et al. 2011). The active site of the PIN domain contains three or four conserved acidic residues that are involved in binding divalent metal ions and catalysis. Structural alignment of Utp23 with SMG6 shows that two corresponding active site residues (Asp31, Asp123) are conserved in Utp23, whereas the other two (Ala60 and Ala100) are mutated (Fig. 2F). The equivalent residue of Ala100 has been shown to be essential for nuclease activity of SMG6 (Glavan et al. 2006) and for the function of Utp24 and Nob1 (Fatica et al. 2004; Bleichert et al. 2006). Moreover, the lethal phenotype of *utp23Δ* strain can be fully rescued by the D31N or D123N mutant of Utp23, indicating that the two conserved active site residues are functionally dispensable (Fig. 3C). This agrees with the previous complementation study using a conditional tetracycline repressible *utp23* strain (Bleichert et al. 2006). Glu101 and Glu149 are located in the vicinity of the active site and may act as alternative catalytic residues (Fig. 3A). This is unlikely since their replacement by uncharged glutamine caused no growth phenotype (Fig. 3C). The structural and mutational data strongly implicate that Utp23 contains an inactive PIN nuclease domain.

tionally dispensable (Fig. 3C). This agrees with the previous complementation study using a conditional tetracycline repressible *utp23* strain (Bleichert et al. 2006). Glu101 and Glu149 are located in the vicinity of the active site and may act as alternative catalytic residues (Fig. 3A). This is unlikely since their replacement by uncharged glutamine caused no growth phenotype (Fig. 3C). The structural and mutational data strongly implicate that Utp23 contains an inactive PIN nuclease domain.

### The helix $\alpha_1$ and C-terminal tail are important functional elements

To gain insight into the potential functional site of Utp23, we mapped conserved residues onto the structure (Fig. 3A). Remarkably, helix  $\alpha_1$  exposes several highly conserved basic residues, namely Arg5, Lys7, Arg10, and Lys11. Lys4 is also moderately conserved as the corresponding position is occupied by lysine in 63% of Utp23 proteins and by threonine or serine in 28% of Utp23 proteins. The abundant basic residues render helix  $\alpha_1$  the most positively charged region in the electrostatic potential surface of Utp23 (Fig. 3B). Moreover, a few conserved residues are scattering in the neighborhood of helix  $\alpha_1$ , including Arg23 and Pro25 on one side and Val147, Glu149, and Arg131 on the other side.

Next, we asked the functional importance of these conserved surface residues. To accomplish this, we constructed a *UTP23* shuffle strain in which the genomic copy of *UTP23* gene was deleted and replaced by *UTP23* gene on a



**FIGURE 2.** Structure of Utp23 PIN domain. (A) Ribbon representation of the domain-swapped dimer. Two molecules are shown in green and magenta. (B) The  $2fo-fc$  electron density contoured at the  $1.5 \sigma$  level is overlaid on the refined structure. (C) Monomeric structure model of Utp23 PIN domain. Two orthogonal views are shown. The  $\beta$ -strands,  $\alpha$ -helices, and loops are colored in yellow, red, and green, respectively. The zinc ion is shown as a yellow ball. The secondary structures and the N and C termini are labeled. (D) Topology diagram of the Utp23 PIN domain. The secondary structures that are common to other PIN domains are colored green for  $\alpha$ -helices and yellow for  $\beta$ -strands, and those unique to Utp23 are shown as empty boxes. Four zinc-coordinating residues are labeled. (E) Close-up view of the zinc-finger motif. (F) Structural superimposition of Utp23 (green) and SMG6 (PDB code 2HWW, gray). The catalytic residues of SMG6 and their Utp23 equivalents are shown as sticks. Utp23 residues are labeled.

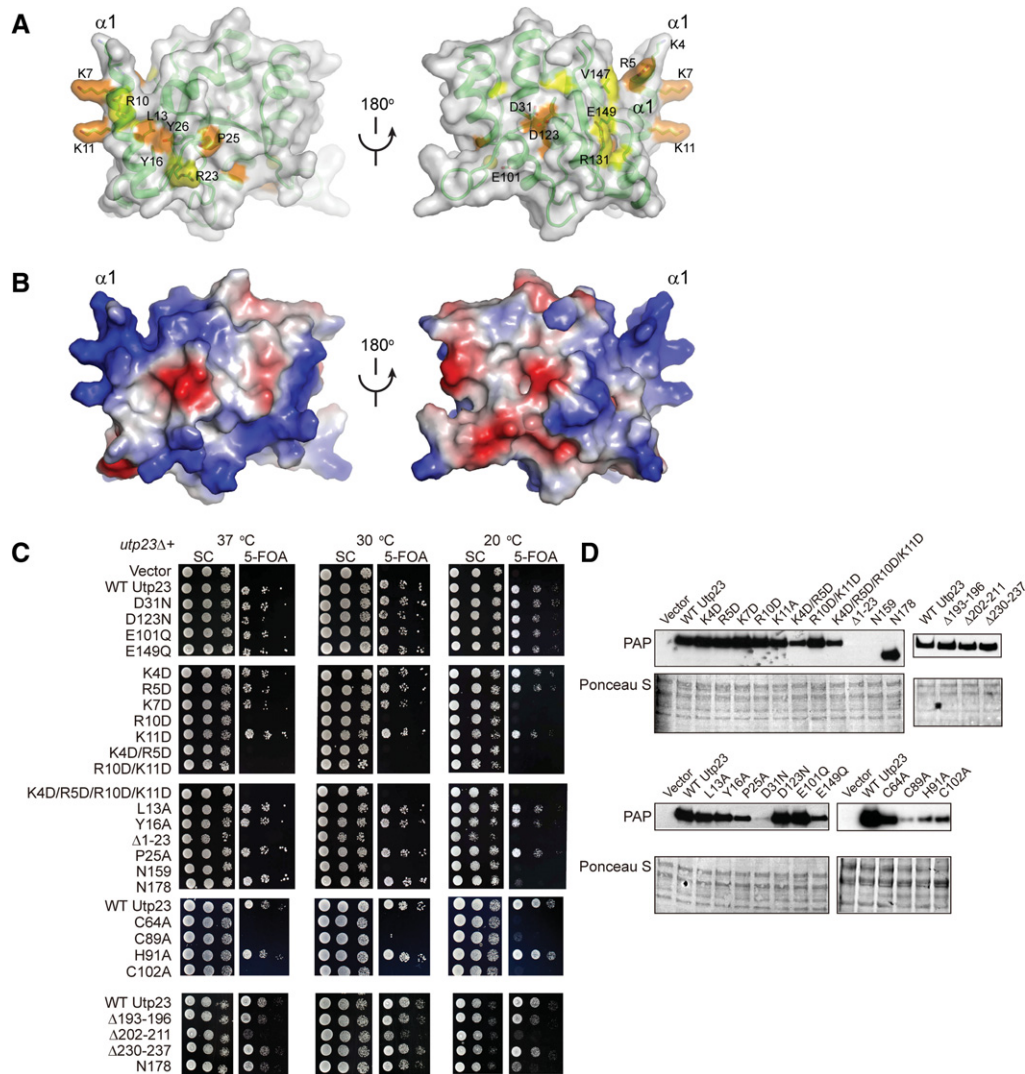
*URA3* plasmid. The shuffle strain was transformed with a *LEU* plasmid expressing TAP-tagged Utp23 variants under the control of the *GPD* promoter. The function of Utp23 variants was examined by growth assay in the presence of 5-FOA, which counter-selects the *URA3 UTP23* plasmid, at temperatures 20°C, 30°C, and 37°C (Fig. 3C). All live strains were recovered from 5-FOA plates and further grown on YPD plates at different temperatures to verify growth phenotype (Supplemental Fig. S2).

The conserved basic residues on helix  $\alpha 1$  were substituted individually or in combination by negatively charged aspartic acid. The K4D and R5D single mutants displayed normal growth, but their double mutant cannot survive at all temperatures, indicating that both Lys4 and Arg5 are functionally important. The K7D mutant grew normally at 30°C and 37°C yet slowly at 20°C, indicating a cold-sensitive phenotype. The R10D mutant was lethal, whereas the K11D mutant had no detectable growth defect. As expected for additive mutational effect, the double mutation R10D/K11D and the quadruple mutation K4D/R5D/R10D/K11D were also lethal. Since these analyzed residues are all located on the exposed surface of helix  $\alpha 1$ , their mutations should minimally affect the stability and folding of the PIN domain structure.

Collectively, these results indicate that helix  $\alpha 1$  is a key functional element of Utp23.

We also analyzed the functional importance of a few other conserved residues by mutagenesis. Residues Leu13 and Tyr16 mediate hydrophobic packing of helix  $\alpha 1$ , and their individual Ala substitutions had negligible effect on cell growth. Alanine replacement of the conserved Pro25 residue was also neutral.

Besides the PIN domain, Utp23 possesses a long tail at the C terminus. The yeast cell expressing only a C-terminal truncation mutant of Utp23 (N178, residues 1–178) showed a cold-sensitive phenotype, indicating that the deleted C-terminal residues are functionally important. The C-terminal tail is composed mainly of low complexity, hydrophilic sequences (Fig. 1), and it also harbors a highly conserved short motif (PNPLSVK) at residues 202–211 and two flanking moderately conserved K/R tracks at residues 193–196 and 230–237. We asked whether these regions are responsible for the observed growth defect in the N178 mutant. The growth assay showed that deletion of residues 202–211 inhibited yeast growth at all three temperatures even more severely than the N178 mutant. Removal of each K/R track caused minor or no effect (Fig. 3C; Supplemental Fig. S2). Hence, the



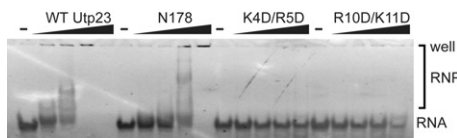
**FIGURE 3.** Functional sites of Utp23. (A) Conservation surface of Utp23 PIN domain. Two opposite orientations are shown. Residues that are conserved in >95% and 95%–80% of 134 Utp23 sequences are colored in orange and yellow, respectively. The surface is semitransparent and overlaid on the ribbon representation of Utp23. Some conserved residues are shown as sticks and labeled. (B) Electrostatic potential surface of Utp23 PIN domain shown in the same orientations as in A. The surface is colored from blue to red for positively to negatively charged regions. (C) Yeast growth assay of Utp23 mutants. The *utp23 $\Delta$*  haploid strain complemented by a *URA3 UTP23* plasmid was transformed with *LEU2* plasmids expressing GPD-driven *UTP23*, no *UTP23* (Vector), or the indicated *utp23* mutations. Tenfold serial dilutions of the transformants were spotted onto plates containing SC medium or SC with 5-FOA to counter-select for the *URA3 UTP23* plasmid and incubated at 37°C, 30°C, or 20°C. (D) Expression of TAP-tagged Utp23 and its mutants from plasmid. Equal amounts of total protein were analyzed by SDS-PAGE, Ponceau S staining, and Western blotting using PAP to detect Utp23-TAP.

conserved residues 202–211 constitute the key functional site in the C-terminal tail.

Deletion of residues 1–23 ( $\Delta 1-23$ ) or residues 160–254 (N159) caused lethality, but this may be due to no expression of the two mutant proteins (Fig. 3D). The expression level of the D31N mutant was low but apparently sufficient for supporting growth. All other mutant proteins were well expressed; hence, their growth phenotype should be ascribed to their impaired function rather than any problem in protein expression. In summary, our mutagenesis and yeast complementation analyses reveal that helix  $\alpha 1$  and the C-terminal tail contain key functional residues.

### Utp23 is an RNA-binding protein

The presence of degenerate PIN ribonuclease domain and the essential function of basic residues in helix  $\alpha 1$  suggest that Utp23 may bind RNA. Indeed, electrophoretic mobility shift assay (EMSA) showed that Utp23 was capable of forming complexes with a fragment of 18S rRNA (Fig. 4). Excessive protein also led to the formation of very large complexes that stuck in the well, indicative of nonspecific binding. The binding affinity was estimated to be in the low  $\mu\text{M}$  range based on the concentrations of RNA and protein used.



**FIGURE 4.** RNA binding assay of Utp23 and its mutants. Ten picomoles of a 18S rRNA fragment (nt 745–859) was incubated with 0, 10, 20, 40, and 100 pmol of Utp23 proteins, resolved in a native gel and stained by GelRed.

To locate the RNA-binding residue of Utp23, recombinant mutant proteins were prepared and assessed. The C-terminal truncation mutant N178 showed a reduced RNA-binding activity, whereas two helix  $\alpha 1$  mutants (K4D/R5D and R10D/K11D) completely lost RNA binding. Similar results were obtained for other RNAs with various sequences and structures (data not shown). These results indicate that Utp23 possesses at least a general RNA-binding activity, and helix  $\alpha 1$  is a major RNA-binding structural element.

### The C-terminal tail of Utp23 is involved in binding snR30

Utp23 is a component of 90S pre-ribosomes and specifically associates with snR30 (Bleichert et al. 2006; Hoareau-Aveilla et al. 2011). We asked whether the functionally impaired Utp23 mutants identified above are defective in binding snR30 and 90S pre-ribosomes. To this end, TAP-tagged Utp23 expressed in the *UTP23* shuffle strain was subjected to immunoprecipitation (IP) with IgG Sepharose, and the coprecipitated RNA was analyzed by Northern blot. IP of wild-type Utp23-TAP showed that snR30 was significantly enriched over U3, snR10, and U14, which are three other snoRNAs involved in early 18S processing (Fig. 5, lane 9). Two control RNAs, U1 snRNA and Trp tRNA, were hardly detectable, confirming the specificity of our IP experiments. This result agrees with the previous finding that Utp23 specifically associates with snR30 (Hoareau-Aveilla et al. 2011).

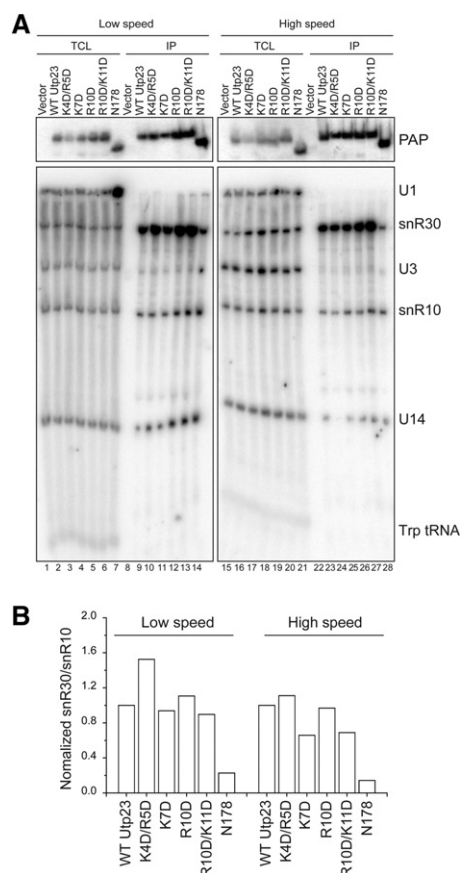
When the cell lysate was subjected to an additional ultracentrifugation step to remove large RNPs including 90S particles, the profile of coprecipitated RNAs was not changed (Fig. 5). The detected association of Utp23 with snR30 apparently occurred independently of 90S pre-ribosome formation. The level of coprecipitated U3, which is a stable component of 90S pre-ribosomes, was low regardless of ultracentrifugation treatment. This indicates that the interaction of Utp23 with 90S pre-ribosome was weak or transitive and hardly detected in our IP experiments.

The helix  $\alpha 1$  mutants K4D/R5D, K7D, R10D and R10D/K11D coprecipitated snR30 as efficiently as wild-type Utp23 (Fig. 5A,B). In contrast, the C-terminal truncation mutant N178 coprecipitated a much smaller amount of snR30 compared to wild-type Utp23 and the helix  $\alpha 1$  mutants (Fig. 5,

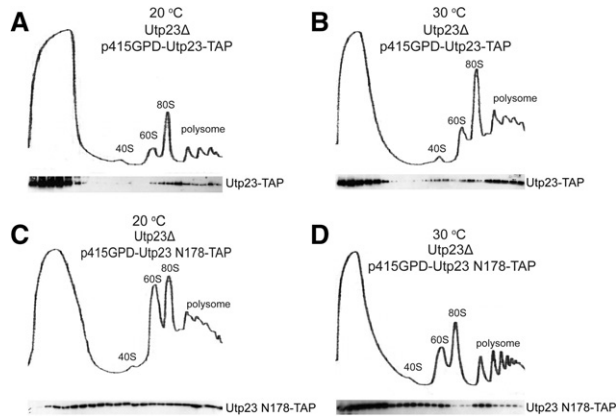
lane 14). This finding demonstrates that the C-terminal tail plays a crucial role in binding snR30.

We found that an increased amount of U3 was coprecipitated by the N178 mutant, but not after the ultracentrifugation treatment to remove large RNPs (Fig. 5, lanes 14,28). The N178 mutant may associate with 90S pre-ribosomes more tightly or for a longer duration than wild-type Utp23. However, association of the N178 mutant with 90S pre-ribosomes appears to be sensitive to experimental variations and was not observed in every experiment (Supplemental Fig. S3).

Since the N178 mutant has a cold-sensitive phenotype, we wondered whether its interaction with snR30 is affected by growth temperature. We grew yeast at 30°C and 20°C and analyzed the interaction between Utp23 and snR30 by IP. The



**FIGURE 5.** Association of Utp23 with snR30. (A) The *UTP23* shuffle strain transformed with an empty *LEU* plasmid (Vector), or *LEU* plasmids expressing TAP-tagged Utp23 or indicated mutants were grown at 30°C. Cell extracts were centrifuged at 10,000g for 30 min (low speed) and additionally at 180,000g for 45 min to deplete large RNPs (high speed). Total cell lysates (TCL, 5%) and immunoprecipitations (IP) of IgG Sepharose were analyzed by Western blotting to detect Utp23-TAP and by Northern blotting with appropriate DNA probes to detect the indicated RNAs. (B) Quantification of snR30. snR10 is used as the internal loading reference because its amount appears to be consistent among different samples. The snR30 to snR10 ratio of Utp23 mutants is further normalized against that of wild-type Utp23 in the same treatment group.



**FIGURE 6.** Sedimentation behavior of Utp23 and the N178 mutant. The *utp23Δ* strains expressing Utp23-TAP (A,B) or its N178 mutant (C,D) from a GPD-driven plasmid were grown at 30°C (B,D) or 20°C (A,C). The cell extracts were fractionated on 7%–50% sucrose gradients, and Utp23-TAP was detected using PAP. The ribosome profiles were recorded at 254 nm.

amount of snR30 coprecipitated by the N178 mutant was reduced to a similar degree at two temperatures (Supplemental Fig. S3), indicating that growth temperature does not directly affect the interaction between the N178 mutant and snR30. It should be noted that the yeast strains used for IP also expressed wild-type Utp23 to support growth of those lethal mutants. The wild-type Utp23 would rescue the cold-sensitive phenotype of the N178 mutant and might mask any temperature-dependent effect on snR30 interaction. However, use of strains without the wild-type Utp23 yielded the same result (Supplemental Fig. S3).

### The C-terminal truncation perturbs pre-ribosomal association of Utp23

To examine whether the C-terminal truncation affects association of Utp23 with pre-ribosomes, we analyzed sedimentation behavior of Utp23 and the N178 mutant at two growth temperatures (Fig. 6). A large fraction of Utp23 was found in the upper part of the gradient as free or snR30-bound proteins (Fig. 6A,B). This was not observed in the previous study (Hoareau-Aveilla et al. 2011), probably because Utp23 was previously expressed under its endogenous promoter, but was overexpressed under the GPD promoter in our strains.

As observed previously, wild-type Utp23 cosedimented with large 90S pre-ribosomes at the 90S and polysome regions, and a small amount of Utp23 can also be detected at the 40S and 60S regions particularly at 30°C (Fig. 6A,B). Compared with wild-type Utp23, the N178 mutant was significantly enriched at the 40S and 60S regions (Fig. 6D). The sedimentation of the N178 mutant was more perturbed at lower temperature 20°C and displayed a near even distribution pattern along the gradient (Fig. 6C). These data suggest

that the C-terminal truncation affects the dynamic association of Utp23 with pre-ribosomes.

## DISCUSSION

A large number of proteins have been identified to be involved in the synthesis of small ribosomal subunits in yeast, yet their structure and functional role have not been characterized in most cases. In this study, we have conducted a detailed structure–function analysis of Utp23, an evolutionarily conserved protein in 90S particles. The crystal structure reveals that the PIN domain of Utp23 shares a core fold with other PIN domains, confirming the previous prediction based on sequence analysis. We also confirm by complementation assay that the putative active residues of Utp23 are dispensable for function (Bleichert et al. 2006), implying that Utp23 is a structural protein rather than an enzyme.

The structure reveals several unique structural elements that are not present in other structurally characterized PIN domains. We identify a novel CCHC Zn finger located at the periphery of the PIN core structure and show that it is essential for Utp23 structure or function. We found that the C64A mutant protein exclusively formed inclusion bodies in *Escherichia coli* and cannot be purified (data not shown), suggesting that the Zn-finger motif is important for maintaining proper structure of Utp23. In addition, other than the four Zn-coordinating ligands, the exposed residues around the Zn finger are not conserved (Fig. 3A), arguing against that the Zn finger is directly involved in molecular interaction.

To the best of our knowledge, this is the first time a Zn-finger motif has been found within a PIN domain. Notably, the four zinc-binding residues are also conserved in Utp24, which is likely an active PIN domain endonuclease for A1 and/or A2 cleavage of pre-rRNA (Bleichert et al. 2006). Despite having different activities and roles in early processing of 18S rRNA, Utp23 and Utp24 share the characteristic zinc-binding residues, suggesting that they are more related to each other with respect to structure and evolution than to other PIN domain proteins.

Another unique structural element of Utp23 is its N-terminal helix  $\alpha$ 1, which exposes many highly conserved basic residues. We show that helix  $\alpha$ 1 is a key functional element of Utp23 since mutations in its exposed basic residues significantly impaired yeast growth. We also found that the basic residues are required for in vitro RNA-binding activity of Utp23, suggesting that helix  $\alpha$ 1 is involved in binding some endogenous RNA targets. Given that helix  $\alpha$ 1 mutants are not defective in binding snR30 RNA, we can speculate that helix  $\alpha$ 1 may be involved in binding rRNA within 90S pre-ribosomes.

Utp23 contains a long and likely unstructured tail at the C terminus of the PIN domain. Our complementation results show that the C-terminal tail is important for yeast growth and harbors a major functional site in conserved residues

202–211. Interestingly, the N178 mutant is defective in binding snR30, which provides mechanistic insight into the function of the C-terminal tail. The growth phenotype of the N178 mutant also demonstrates the functional role of Utp23–snR30 interaction. How the C-terminal tail mediates snR30 binding remains an open question. The C-terminal tail could directly bind snR30 RNA or bind other proteins associated with snR30.

Our gradient sedimentation results reveal that the N178 mutant is more strongly associated with particles migrating at the 40S and 60S regions compared with wild-type Utp23. These particles may represent mature 40S and 60S ribosomal subunits with which Utp23 nonspecifically associates or other uncharacterized pre-ribosomal particles. Understanding the nature of these Utp23-associated particles may shed further light on the ribosome assembly pathway.

In summary, we have revealed the structure and major functional sites of Utp23. These functionally important regions, including the zinc finger, helix  $\alpha$ 1, and the C-terminal tail, are all located outside the core structure of the PIN domain. The ancestral PIN domain of Utp23 apparently evolved new functionality by developing extra structural elements. The Utp23 mutants with different phenotypes identified in this study can be used as tool to further study the assembly and dynamics of 90S pre-ribosomes.

## MATERIALS AND METHODS

### Gene cloning and protein purification

The DNA sequence of Utp23 was PCR-amplified from yeast genomic DNA and cloned into a modified pETDuet-1 (Novagen) vector to express Utp23 protein fused to an N-terminal His<sub>6</sub>-tag and a PreScission cleavage site. The C-terminal truncation mutants were constructed by introducing the stop code at appropriate positions of Utp23 gene with the QuikChange method (Stratagene). *E. coli* BL21 (DE3) cells were transformed and grown in LB broth containing 100  $\mu$ g/mL of ampicillin at 37°C to an OD<sub>600</sub> of 0.8. The protein expression was then induced with 0.2 mM isopropyl-D-1-thiogalactopyranoside for 16 h at 16°C. Cells were pelleted, resuspended in buffer A (50 mM Tris-HCl, pH 7.5, 300 mM NaCl, 5% glycerol) containing 30 mM imidazole and 0.1 mM phenylmethane sulphonylfluoride, and then lysed in a high-pressure JN-3000 PLUS cell disruptor (JNBIO). After clarification by centrifugation, the supernatant was loaded onto a 5-mL HisTrap column (GE Healthcare). The protein was eluted with a linear gradient of 30 to 500 mM imidazole in buffer A. The target protein was pooled and digested with PreScission overnight at 4°C to remove the N-terminal His<sub>6</sub>-tag. The target protein was further purified through a heparin column (GE Healthcare) followed by gel filtration (Superdex 75 10/300 GL, GE Healthcare) pre-equilibrated in 10 mM Tris-HCl, pH 7.5, 200 mM NaCl.

### Crystallization and structure determination

Utp23 1–159 was crystallized at 20°C using the sitting-drop vapor diffusion method by mixing 1  $\mu$ L of protein solution (30 mg/mL

protein, 10 mM Tris-HCl, pH 7.5, 200 mM NaCl) and 1  $\mu$ L of reservoir solution (0.1 M Bis-Tris pH 6.5 or 0.1 M HEPES pH 7.5, 2% v/v tacsimate pH 6.0, 20% w/v polyethylene glycol 3350). The crystals were directly frozen in liquid nitrogen without additional cryoprotection.

Diffraction data were collected at beamline BL17U of the Shanghai Synchrotron Radiation Facility (SSRF) and processed with HKL2000 (Otwinowski and Minor 1997). A 2.9-Å data set was collected at the Zn peak absorbance wavelength and used for SAD phasing in SHARP (Vonnrhein et al. 2007). Following density modification, the experimental electron density map was of sufficient quality to allow manual model building in COOT (Emsley and Cowtan 2004). The structure was refined in Refmac and PHENIX (Murshudov et al. 1999; Adams et al. 2010). The final model was refined against a 2.5-Å data set collected from another crystal. The final model contains two Utp23 molecules with residues 3–159 and residues 3–92 and 97–159, respectively; two zinc ions; and 21 water molecules. The error of coordinates is 0.40 Å based on maximum-likelihood estimation. RAMPAGE analysis showed that 96.1% of the residues are in favored regions and 3.9% are in allowed regions. The structural figures were prepared using PyMOL (DeLano 2002).

### Yeast strains, mediums, and plasmids

Standard protocols were followed for yeast manipulation. Yeast cells were cultured in YPD (1% yeast extract, 2% peptone, 0.003% adenine, and 2% glucose), Synthetic Complete (SC) medium, and appropriate SC dropout mediums (Clontech). The used strains, oligonucleotides, and plasmids are listed in Supplemental Tables S1–S3.

The *UTP23* gene, including the flanking promoter and terminator sequences were PCR-amplified from yeast genomic DNA and cloned into the *URA3* pRS416 plasmid with the In-fusion method (TaKaRa). The plasmid p415GPD-Utp23-TAP that expresses C-terminally TAP-tagged Utp23 under the GPD promoter was kindly provided by Dr. Susan Baserga (Bleichert et al. 2006) and mutated with the QuikChange method. To generate the *UTP23* shuffle strain, the heterozygous deletion diploid *utp23 $\Delta$ /UTP23* (Euroscarf) was transformed with the *URA3 UTP23* plasmid. The transformants were sporulated, and the spores were isolated and germinated. The *utp23 $\Delta$*  haploid strain complemented with the *URA3 UTP23* plasmid was selected in Ura-deficient SC medium containing antibiotic G418 and verified by PCR.

### Yeast growth assay

The *utp23 $\Delta$*  haploid strain carrying pRS416-UTP23 was transformed with p415GPD-Utp23-TAP encoding wild-type or mutant Utp23-TAP. The transformants were inoculated into 2 mL of Ura- and Leu-deficient SC liquid medium and grown at 30°C until the OD<sub>600</sub> value reached 0.6–1.0. Cells were adjusted to an OD<sub>600</sub> of 0.6 and serially diluted 10-fold with sterile water. The dilutions were spotted on SC plates with or without 0.1% 5-fluoroorotic acid (5-FOA) and incubated for 4 d at 37°C and 30°C and for 5 d at 20°C. The live strains in 5-FOA plates were recovered, cultured in liquid YPD medium, spotted onto YPD plates, and grown for 2 d at 37°C and 30°C and for 4 d at 20°C.



## Immunoprecipitation

Yeast *utp23Δ* strains carrying pRS416-Utp23 and p415GPD-Utp23-TAP or mutants were cultured in 10 mL of Leu- and Ura-deficient SC medium for ~16 h and further grown in 1 L of YPD to OD<sub>600</sub> between 0.8 and 1.0. IP was performed as previously described (Hoareau-Aveilla et al. 2011). Cells were collected and resuspended in lysis buffer (20 mM Tris-HCl, pH 8.0, 5 mM Mg-acetate, 200 mM NaCl, 0.2% Triton X-100) supplemented with EDTA-free protease inhibitor cocktail (Roche), 0.5 units/μL RNasin (Promega), and 1 mM DTT. After cells were broken by a high-pressure cell disruptor, lysates were clarified by centrifugation at 10,000g for 30 min. To obtain high speed extracts, half the supernatant was centrifuged at 180,000g for 45 min. IgG Sepharose beads (100 μL) were incubated with 150 OD<sub>260</sub> units of supernatant for 2 h and washed seven times with 1 mL lysis buffer.

For protein analysis, the precipitants were separated on 4%–20% gradient SDS-PAGE gels and immunoblotted with peroxidase anti-peroxidase (1:5000, Sigma). For RNA analysis, RNA was isolated using the hot phenol method, separated in 6% polyacrylamide-8M urea gels, and transferred to Hybond N<sup>+</sup> membranes (GE Healthcare). The following oligonucleotides were used for northern hybridization: U1: 5'-CTGATATCTTAAGGTAAGTAT; snR30: 5'-ATGTCTGCAGTATGGTTTAC; U3: 5'-GGATTGCGGACCAAGCTAA; snR10: 5'-GTGTTACGAATGGCTGTTA; U14: 5'-TCACTCAGA CATCCTAGG; Trp tRNA: 5'-CATTACGAGTGCATGCCTTAC. Northern blotting was conducted as described (Lin et al. 2013).

## Sucrose gradient sedimentation

Sucrose gradient sedimentation analysis was performed as described (Lin et al. 2013). Yeast cells (250–300 mL) were grown to OD<sub>600</sub> = 0.8–1.0 at 30°C. For low temperature treatment, yeast cells were grown at 30°C to an OD<sub>600</sub> of 0.6 and then for 3 h at 20°C.

## EMSA

RNAs were in vitro transcribed by T7 polymerase using PCR products as template and gel purified. RNAs (10 pmol) were assembled with Utp23 or mutant proteins in different molar ratios in 10 μL of binding buffer containing 25 mM HEPES-K (pH 7.6), 100 mM NaCl, 2 mM MgCl<sub>2</sub>, 1 mM DTT, 0.01% NP-40, and 10% glycerol. RNAs were separated in 5% native acrylamide gels and stained with GelRed.

## DATA DEPOSITION

The atomic coordinates and experimental data have been deposited in the Protein Data Bank under accession code 4MJ7.

## SUPPLEMENTAL MATERIAL

Supplemental material is available for this article.

## ACKNOWLEDGMENTS

We thank the staff at the Shanghai Synchrotron Radiation Facility beamline BL17U for help in diffraction data collection. We are grateful to Dr. Susan Baserga (Yale University) for plasmids and

strains. This research was supported by the National Basic Research Program of China (973 Program, 2010CB835402) and the Beijing Municipal Government.

Received June 23, 2013; accepted September 24, 2013.

## REFERENCES

- Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung LW, Kapral GJ, Grosse-Kunstleve RW, et al. 2010. PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66**: 213–221.
- Arcus VL, McKenzie JL, Robson J, Cook GM. 2011. The PIN-domain ribonucleases and the prokaryotic VapBC toxin-antitoxin array. *Protein Eng Des Sel* **24**: 33–40.
- Atzorn V, Fragapane P, Kiss T. 2004. U17/snR30 is a ubiquitous snoRNA with two conserved sequence motifs essential for 18S rRNA production. *Mol Cell Biol* **24**: 1769–1778.
- Bleichert F, Granneman S, Osheim YN, Beyer AL, Baserga SJ. 2006. The PINc domain protein Utp24, a putative nuclease, is required for the early cleavage steps in 18S rRNA maturation. *Proc Natl Acad Sci* **103**: 9464–9469.
- Clissold PM, Ponting CP. 2000. PIN domains in nonsense-mediated mRNA decay and RNAi. *Curr Biol* **10**: R888–R890.
- DeLano WL. 2002. *The PyMOL user's manual*. DeLano Scientific, San Carlos, CA.
- Dragon F, Gallagher JE, Compagnone-Post PA, Mitchell BM, Porwancher KA, Wehner KA, Wormsley S, Settlege RE, Shabanowitz J, Osheim Y, et al. 2002. A large nucleolar U3 ribonucleoprotein required for 18S ribosomal RNA biogenesis. *Nature* **417**: 967–970.
- Emsley P, Cowtan K. 2004. *Coot*: Model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**: 2126–2132.
- Fatica A, Oeffinger M, Dlakić M, Tollervy D. 2003. Nob1p is required for cleavage of the 3' end of 18S rRNA. *Mol Cell Biol* **23**: 1798–1807.
- Fatica A, Tollervy D, Dlakić M. 2004. PIN domain of Nob1p is required for D-site cleavage in 20S pre-rRNA. *RNA* **10**: 1698–1701.
- Fayet-Lebaron E, Atzorn V, Henry Y, Kiss T. 2009. 18S rRNA processing requires base pairings of snR30 H/ACA snoRNA to eukaryote-specific 18S sequences. *EMBO J* **28**: 1260–1270.
- Glavan F, Behm-Ansmant I, Izaurralde E, Conti E. 2006. Structures of the PIN domains of SMG6 and SMG5 reveal a nuclease within the mRNA surveillance complex. *EMBO J* **25**: 5117–5125.
- Grandi P, Rybin V, Bassler J, Petfalski E, Strauss D, Marzioch M, Schafer T, Kuster B, Tschochner H, Tollervy D, et al. 2002. 90S pre-ribosomes include the 35S pre-rRNA, the U3 snoRNP, and 40S subunit processing factors but predominantly lack 60S synthesis factors. *Mol Cell* **10**: 105–115.
- Henras AK, Soudet J, Geras M, Lebaron S, Caizergues-Ferrer M, Mougin A, Henry Y. 2008. The post-transcriptional steps of eukaryotic ribosome biogenesis. *Cell Mol Life Sci* **65**: 2334–2359.
- Hoareau-Aveilla C, Fayet-Lebaron E, Jady BE, Henras AK, Kiss T. 2011. Utp23p is required for dissociation of snR30 small nucleolar RNP from preribosomal particles. *Nucleic Acids Res* **40**: 3641–3652.
- Holm L, Sander C. 1993. Protein structure comparison by alignment of distance matrices. *J Mol Biol* **233**: 123–138.
- Karstein K. 2011. Inside the 40S ribosome assembly machinery. *Curr Opin Chem Biol* **15**: 657–663.
- Kiss T, Fayet-Lebaron E, Jady BE. 2010. Box H/ACA small ribonucleoproteins. *Mol Cell* **37**: 597–606.
- Kressler D, Hurt E, Bassler J. 2010. Driving ribosome assembly. *Biochim Biophys Acta* **1803**: 673–683.
- Levin I, Schwarzenbacher R, Page R, Abdubek P, Ambing E, Biorac T, Brinen LS, Campbell J, Canaves JM, Chiu HJ, et al. 2004. Crystal structure of a PIN (PiIT N-terminus) domain (AF0591) from *Archaeoglobus fulgidus* at 1.90 Å resolution. *Proteins* **56**: 404–408.

- Lin J, Lu J, Feng Y, Sun M, Ye K. 2013. An RNA-binding complex involved in ribosome biogenesis contains a protein with homology to tRNA CCA-adding enzyme. *PLoS Biol* **11**: e1001669. doi: 1001610.1001371/journal.pbio.1001669.
- Morrissey JP, Tollervey D. 1993. Yeast snR30 is a small nucleolar RNA required for 18S rRNA synthesis. *Mol Cell Biol* **13**: 2469–2477.
- 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.
- Osheim YN, French SL, Keck KM, Champion EA, Spasov K, Dragon F, Baserga SJ, Beyer AL. 2004. Pre-18S ribosomal RNA is structurally compacted into the SSU processome prior to being cleaved from nascent transcripts in *Saccharomyces cerevisiae*. *Mol Cell* **16**: 943–954.
- Otwinowski Z, Minor W. 1997. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol* **276**: 307–326.
- Panse VG, Johnson AW. 2010. Maturation of eukaryotic ribosomes: Acquisition of functionality. *Trends Biochem Sci* **35**: 260–266.
- Pertschy B, Schneider C, Gnädig M, Schäfer T, Tollervey D, Hurt E. 2009. RNA helicase Prp43 and its co-factor Pfa1 promote 20 to 18 S rRNA processing catalyzed by the endonuclease Nob1. *J Biol Chem* **284**: 35079–35091.
- Phipps KR, Charette JM, Baserga SJ. 2011. The SSU processome in ribosome biogenesis—progress and prospects. *Wiley Interdiscip Rev RNA* **2**: 1–21.
- Stavridi ES, Huyen Y, Loreto IR, Scolnick DM, Halazonetis TD, Pavletich NP, Jeffrey PD. 2002. Crystal structure of the FHA domain of the Chfr mitotic checkpoint protein and its complex with tungstate. *Structure* **10**: 891–899.
- Tschochner H, Hurt E. 2003. Pre-ribosomes on the road from the nucleolus to the cytoplasm. *Trends Cell Biol* **13**: 255–263.
- Vonrhein C, Blanc E, Roversi P, Bricogne G. 2007. Automated structure solution with autoSHARP. *Methods Mol Biol* **364**: 215–230.
- Watkins NJ, Bohnsack MT. 2011. The box C/D and H/ACA snoRNPs: Key players in the modification, processing and the dynamic folding of ribosomal RNA. *Wiley Interdiscip Rev RNA* **3**: 397–414.