Interaction between Ribosome Assembly Factors Krr1 and Faf1 Is Essential for Formation of Small Ribosomal Subunit in Yeast*

Received for publication, May 27, 2014, and in revised form, June 24, 2014 Published, JBC Papers in Press, July 2, 2014, DOI 10.1074/jbc.M114.584490

Sanduo Zheng (郑三多),†‡§¶†§‡, Pengfei Lan (兰鹏飞),‡§¶1, Ximing Liu (刘希明),† and Keqiong Ye (叶克穷)‡§¶1

From the †Department of Biochemistry and Molecular Biology, College of Life Sciences, Beijing Normal University, Beijing 100875, the ‡Graduate School of Peking Union Medical College and Chinese Academy of Medical Sciences, Beijing 100730, the §National Institute of Biological Sciences at Beijing, Beijing 102206, and the ¶Laboratory of RNA Biology, Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, China

Background: The 90 S precursor of small ribosomal subunits is an enormous and poorly understood structure.

Results: A subcomplex structure of 90 S proteins Krr1 and Faf1 was determined.

Conclusion: The KH domains of Krr1 are versatile in protein interaction, and the Krr1-Faf1 interaction is essential for forming functional 90 S pre-ribosomes.

Significance: The study provides structural and functional insight of a protein-protein interaction within 90 S pre-ribosomes.

Ribosome formation in Saccharomyces cerevisiae requires a large number of transiently associated assembly factors that coordinate processing and folding of pre-rRNA and binding of ribosomal proteins. Krr1 and Faf1 are two interacting proteins present in early 90 S precursor particles of the small ribosomal subunit. Here, we determined a co-crystal structure of the core domain of Krr1 bound to a 19-residue fragment of Faf1 at 2.8 Å resolution. The structure reveals that Krr1 consists of two packed K homology (KH) domains, KH1 and KH2, and resembles archaeal Dim2-like proteins. We show that KH1 is a divergent KH domain that lacks the RNA-binding GXXG motif and is involved in binding another assembly factor, Kri1. KH2 contains a canonical RNA-binding surface and additionally associates with an α-helix of Faf1. Specific disruption of the Krr1-Faf1 interaction impaired early 18 S rRNA processing at sites A0, A1, and A2 and caused cell lethality, but it did not prevent incorporation of the two proteins into pre-ribosomes. The Krr1-Faf1 interaction likely maintains a critical conformation of 90 S pre-ribosomes required for pre-rRNA processing. Our results illustrate the versatility of KH domains in protein interaction and provide insight into the role of Krr1-Faf1 interaction in ribosome biogenesis.

The ribosome of yeast Saccharomyces cerevisiae is assembled from four rRNAs and 79 ribosomal proteins (r-proteins) through a complex and highly dynamic process (1–5). This process begins in the nucleolus, where rRNA repeats are transcribed by RNA polymerase I into a long 35 S precursor rRNA (pre-rRNA). The 35 S pre-rRNA undergoes extensive modification and nucleolytic processing to produce mature 18 S, 5.8 S, and 25 S rRNAs. The 5 S rRNA is transcribed separately by RNA polymerase III. Numerous small nuclear RNAs (snoRNAs) function as guides to direct site-specific modification of rRNA. A few snoRNAs, such as the conserved U3, U14, and snR30/U17, are involved in rRNA processing. Genetic and biochemical studies have identified ~200 protein assembly factors in yeast required for ribosome synthesis. These factors include enzymes, such as RNA helicases, AAA+ ATPases, GTPases, kinases, nucleases, and RNA modification enzymes, many proteins with known protein- or RNA-binding domains, and additional proteins containing no recognizable domain. Ribosome assembly factors transiently associate with pre-rRNA at specific stages, forming distinct pre-ribosomal particles, but the molecular function of most assembly factors is not understood.

The earliest pre-ribosome, often termed the 90 S pre-ribosome or the small subunit processome, is assembled cotranscriptionally on the nascent pre-rRNA transcript and can be visualized as a terminal ball by electron microscopy of rDNA spreads (6, 7). In addition to the 35 S pre-rRNA, the 90 S particle consists of nearly 50 assembly factors, U3 snoRNA, and a subset of small subunit r-proteins (1, 8, 9). The U3 snoRNA forms multiple base-pairing interactions with the 5′-external transcribed spacer and 18 S regions of pre-rRNA, and it is essential for the formation of 90 S pre-ribosomes (1). Within the 90 S pre-ribosome, the pre-rRNA is cleaved at sites A0, A1, and A2, leading to excision of the 5′-external transcribed spacer and separation of the 20 S and 27 SA2 pre-rRNAs; these pre-rRNAs are destined to the small 40 S and large 60 S ribosomal subunit, respectively. The 20 S pre-rRNA, packed in pre-40 S particles, is exported to the cytoplasm and matures into 18 S particles.

*This work was supported by National Natural Science Foundation of China Grant 3125007, Strategic Priority Research Program of the Chinese Academy of Sciences Grant XDB08010203, Ministry of Science and Technology of China (National Basic Research Program of China) Grant 2010CB835402, and the Beijing Municipal Government.

The atomic coordinates and structure factors (code 4QMF) have been deposited in the Protein Data Bank (http://wwwpdb.org/).

† Both authors contributed equally to this work.
‡ To whom correspondence should be addressed. Tel.: 86-10-80726688, ext. 8550; Fax: 86-10-80728592; E-mail: yekeqiong@nibs.ac.cn.
§ The abbreviations used are: r-protein, ribosomal protein; KH, K-homology; KH1, KH domain 1; KH2, KH domain 2; pre-rRNA, precursor RNA; snoRNA, small nuclear RNA; DLL, D152A, L155D, and L159D triple mutation of glyceraldehyde-3-phosphate dehydrogenase.
rRNA after cleavage at site D (10). Compared with 90 S particles, pre-40 S particles have a much simpler composition with most of 40 S r-proteins and a handful of assembly factors (10, 11).

Understanding the structure and assembly of the enormous 90 S particle represents a major challenge for the field. One step toward this goal is to elucidate how each protein interacts with other protein and RNA components using methods with the highest attainable resolution. In this regard, a few proteins of the 90 S particle were shown to form independent subcomplexes, including UTPA (also known as uUTP), UTPB, UTPC, U3 snoRNP, the Mpp10-Imp3-Imp4 complex, and the Bms1-Rcl1 complex (8, 12–17). Analysis of assembly interdependence has revealed a hierarchical assembly order for several subcomplexes (18, 19). The interaction networks among seven proteins of UTPA and among six proteins of UTPB have been mapped by yeast two-hybrid assays (20, 21). Most recently, protein cross-linking and mass spectrometry analysis have mapped spatial closeness of UTPB components at single residue resolution (22). The RNA-protein UV cross-linking approach has been used to locate the precise RNA-binding site of 40 S assembly factors (23–26). However, high resolution structural information is very limited for the 90 S particle (other pre-ribosomal particles as well) and is available only for the H/ACA and C/D types of snoRNP (27) and a few individual components (28–33).

In this study, we focus on two interacting proteins, Krr1 and Faf1, present in the 90 S particle. Both are essential nucleolar proteins in yeast that are required for 18 S rRNA processing at sites A0, A1, and A2 and for the formation of small ribosomal subunits (34–37). They copurified with other components of the 90 S particle in tandem affinity purification (TAP) experiments (8, 38). Krr1 contains a KH homology (KH) domain, which is one of the most abundant nucleic acid-binding motifs recognizing single-stranded RNA or DNA (39). Besides putative RNA binding activity, Krr1 also binds proteins. It binds Faf1 in a yeast two-hybrid assay (37) and interacts genetically and physically with Kri1, which is another factor essential for 40 S formation (34). Krr1 is universally conserved in eukaryotes, and its mRNA was identified in human, Drosophila, Schizosaccharomyces pombe, and humans have been shown to be involved in ribosome biogenesis (40–42). Faf1 contains no recognizable domain and has homologs found so far only in Ascomycetes.

The assembly of the 40 S ribosome requires another KH domain protein, Dim2/Pno1 (43–45). Dim2 is associated with both the early 90 S particle and the late pre-40 S particle and localizes to the nucleolus and the cytoplasm (8, 10). Both Krr1 and Dim2 contain a single conserved KH domain in archaea. A Dim2-like protein that has tandem KH domains is implicated in ribosome biogenesis and translation initiation. The crystal structure of Pyrococcus horikoshii Dim2 (PhDim2) has been determined in both the free state and in complex with a 3’ end fragment of 16 S rRNA and the translation initiation factor eIF2α (46, 47).

We have mapped the interacting region of Krr1 and Faf1 and determined a co-crystal structure of the Krr1 core domain bound to a short fragment of Faf1. The complex structure reveals the presence of tandem packed KH domains in Krr1 and a novel protein-binding mode of the KH domain. We demonstrated that the Krr1-Faf1 interaction is essential for 18 S rRNA processing and yeast growth. However, disruption of the Krr1-Faf1 interaction did not prevent their incorporation into 90 S particles, suggesting that the interaction is required for maintaining a functional conformation of 90 S particles.

**Experimental Procedures**

**DNA Cloning and Protein Purification**—DNA cloning was mainly performed with the nonligation-based methods In-fusion (TaKaRa) and Transfer-PCR (48). The **KRR1** gene was amplified by PCR from yeast genomic DNA, cloned into a modified pET28a vector (Novagen), and expressed with an N-terminal His6-Smt3 tag. The **FAF1** gene was cloned into the multiple cloning site 1 of a modified pETDuet-1 vector and expressed with an N-terminal His6-GST tag followed by a PreScission cleavage site. Point and deletion mutations were generated by the QuikChange method and confirmed by DNA sequencing.

His6-Smt3-tagged Krr1 was expressed in Escherichia coli Rosetta 2 (DE3) (Novagen). Bacteria were cultured in LB medium at 37 °C to an A600 of 0.8, and protein expression was then induced with 0.5 mM isopropyl 1-thio-β-D-galactopyranoside overnight at 18 °C. Cells were harvested, resuspended in buffer A (50 mM Tris-HCl, pH 8.0, and 500 mM NaCl), and lysed by sonication. After centrifugation, Krr1 was purified with a HisTrap column (GE Healthcare) and eluted with 500 mM imidazole in buffer A. The pooled fractions of Krr1 were incubated with Ulp1 for 1 h on ice to cleave the His6-Smt3 tag. The sample was loaded directly onto a heparin column (GE Healthcare) and eluted by ~700 mM NaCl with a linear salt gradient in 50 mM Tris-HCl, pH 8.0. Krr1(32–222) was labeled with selenomethionine (SeMet) in M9 medium by blocking methionine biosynthesis. The SeMet-labeled protein was purified by the same procedure as the unlabeled protein except that the lysis buffer was supplemented with 1 mM DTT.

His6-GST-tagged Faf1 was expressed in BL21-Gold (DE3) cells. Protein expression, cell lysis, and HisTrap chromatography were performed using the same method as for Krr1. The His6-GST tag of Faf1 was cleaved with PreScission overnight at 4 °C. The sample was diluted 3-fold with 25 mM HEPES-KOH, pH 7.6, and loaded onto a Q column (GE Healthcare). The flow-through containing Faf1 was collected and concentrated with 3-kDa cutoff ultrafiltration devices (Amicon).

For assembly and purification of the Krr1-Faf1 complex, individually purified Krr1 and Faf1 proteins were mixed in a 1:2 molar ratio and incubated on ice for 1 h. The binary complex was separated from excessive free Faf1 on a Superdex S200 10/300 column equilibrated in 10 mM Tris-HCl, pH 8.0, and 300 mM NaCl.

**Crystallization, Data Collection, and Structure Determination**—The native and SeMet-labeled complex of Krr1(32–222) and a Faf1 fragment containing residues 145–169 and 199–220 were crystallized at 20 °C using the hanging-drop vapor diffusion method by mixing 1 μl of protein (20 mg ml⁻¹ in 10 mM Tris-HCl, pH 8.0, and 300 mM NaCl) and 1 μl of reservoir solution containing 0.2 M tri-ammonium citrate and 20% (w/v) PEG 3350. Rod-shaped crystals appeared after 1 day of crystallization.
for native and SeMet-labeled proteins. The crystals were cryo-protected in 20% glycerol made in the reservoir solution and flash-frozen in liquid nitrogen.

Diffraction data were collected at the Shanghai Synchrotron Radiation Facility beamline BL17U and processed by HKL2000 (49). The crystal belongs to space group $P_2_1$ and contains two copies of the Krr1-Faf1 heterodimer per asymmetric unit. The phases were calculated with SHARP using the single-wavelength anomalous dispersion method based on a selenium derivative dataset collected at a wavelength of 0.9793 Å and solvent-modified (50). The model was built in Coot and refined in Refmac and Phenix (51–53). The model-derived phases were iteratively combined with the experimental single-wavelength anomalous dispersion phases in SHARP to improve the electron density map. The final model includes two Krr1 molecules with residues 38–211, two Faf1 molecules with residues 144–163 (Pro-144 is from vector), and 14 water molecules. RAP PAGE analysis showed that 96.9% of the residues are in favorable regions, 2.9% in allowed regions, and 0.3% in outlier regions (54).

**GST Pulldown Assay**—His$_6$-GST-tagged Faf1 and its variants were expressed and purified with Ni$^{2+}$ beads. Krr1(32–222) was purified as described above, and full-length Krr1 was briefly purified with Ni$^{2+}$ beads followed by Ulp1 cleavage. For interaction analysis, individually purified GST-Faf1 and Krr1 proteins were mixed with ~10 μM concentrations in a volume of 200 μL. The mixtures were incubated with 15 μL of glutathione-Sepharose beads and gently rotated for 30 min at 4 °C. After the beads were washed three times with 1 ml of buffer A (50 mM Tris-HCl, pH 8.0, and 500 mM NaCl), the bond protein was eluted with 20 μL of 10 mM glutathione in buffer A. The input and eluate samples of glutathione-Sepharose beads were resolved using SDS-PAGE and stained with Coomassie Blue.

**Yeast Plasmids, Strains, and Medium**—Yeast works were performed according to standard protocols. Strains used in this study are listed in Table 1. Genomic tagging was performed using the one-step PCR strategy. Yeast cells were grown in YPDA (1% yeast extract, 2% peptone, 0.003% adenine, 2% glucose), YPGA (1% yeast extract, 2% peptone, 0.003% adenine, 2% galactose), Synthetic Complete (SC) medium, and appropriate SC dropout medium (Clontech).

The FAF1 gene, including 295-bp sequences upstream of the start codon and 652-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-FAF1. The FAF1 ORF was also cloned into LEU2 p415GPD-HA and HIS3 p413GPD-TAP plasmids to express the N-terminally HA-tagged and C-terminally TAP-tagged Faf1 protein under the control of glyceraldehyde-3-phosphate dehydrogenase (GPD) promoter, respectively. Point and deletion mutations were generated by the QuikChange method with proper primers and confirmed by DNA sequencing.

For construction of the FAF1 shuffle strain, BY4741 was transformed with pRS416-FAF1, and the genomic FAF1 gene was replaced with a kanMX6 cassette from plasmid pFA6a-kanMX6 by homologous recombination (55). Positive clones were selected on Ura-deficient SC medium supplemented with antibiotic G418. Clones were confirmed by PCR with appropriate primers to make sure that recombination occurred in the genome but not in the pRS416-FAF1 plasmid.

The KM804 strain (W303a, faf1::LEU2, pRS314-GAL1-myc-FAF1) was kindly provided by Keiko Mizuta (Hiroshima University, Japan) and was described previously (36). The KRR1 gene in KM804 was tagged with a C-terminal 6× hemagglutinin (HA) epitope using the 6HA-kanMX4 cassette from pYPM-14 vector (56).

**Growth Assay**—The faf1Δ strain containing pRS416-FAF1 was transformed with a p415GPD plasmid expressing wild-type or mutant HA-Faf1. The transformant was grown in 1 ml of Ura- and Leu-deficient SC medium overnight at 30 °C. The culture was adjusted to an $A_{600}$ of 0.2 and 5-fold serially diluted in 96-well plates. Ten microliters of cells were spotted onto SC plates with or without 5-fluoroorotic acid and grown for 3 days at 20, 30, and 37 °C.

**Sucrose Gradient Sedimentation, Western Blot, and Northern Blot**—The KM804/KRR1-HA strain was transformed with an empty p413GPD plasmid or plasmids expressing wild-type or mutant Faf1-TAP. The culture was first grown in YPGA medium and then shifted to YPGA medium.

Sucrose gradient sedimentation, Western blot, and Northern blot were performed as described previously (32). The following antibodies were used with appropriate dilution ratios: anti-HA antibody (1:3000, MBL), anti-Myc antibody (1:3000, MBL), and peroxidase anti-peroxidase (1:5000, Sigma). Oligonucleotides for Northern blot are as follows: a, 5′-CATGGCTT-AATCTTTGAGAC-3′; b, 5′-CTCCGCTTATTGATATGC-3′; c, 5′-GGTTTTAAATTGTCCTA-3′; and d, 5′-TGAAAG-GGAATATGACGCT-3′.

**Yeast Two-hybrid Assay**—Two-hybrid assays were performed with the Matchmaker system (Clontech). Krr1 was cloned into the pGBK7T plasmid encoding the Gal4 DNA-binding domain, and Faf1 or Kir1 was cloned into the pGADT7 plasmid encoding the Gal4 DNA activation domain. The two plasmids were cotransformed into the yeast AH109 strain. Leu$^+$ Trp$^+$ transformants were selected. The transformants were grown in 1 ml of Leu- and Trp-deficient SC medium overnight at 30 °C. 5-Fold serial dilutions of cells (10 μL) were spotted on SC plates lacking Leu, Trp, and His and containing 0, 1, or 5 mM 3-amino-1,2,4-triazole. The plates were incubated at 30 °C for 3 days.

### Table 1

| Strain            | Genotype                             | Source         |
|------------------|--------------------------------------|----------------|
| BY4741           | MATa, his3Δ1, leu2Δ10, met15Δ0, ura3Δ0 | Euroscarf      |
| FAF1 shuffle     | BY4741, faf1Δ::kanMX6, [pRS416-FAF1]| This study     |
| W303a            | MATa, leu2-3, 112, his3Δ11, 15, trp1-1, ura3-1, ade2-1, can1-100 | K. Mizuta     |
| KM804            | W303a, faf1::LEU2, [pRS314-GAL1-myc-FAF1] | K. Mizuta     |
| KM804/KRR1-HA    | W303a, faf1::LEU2, KRR1-6HA::kanMX4[pRS314-GAL1-myc-FAF1] | This study     |

**Structure of Krr1-Faf1 Complex**
RESULTS

Structure Determination of the Krr1 and Faf1 Complex—Krr1 and Faf1 were previously reported to interact with each other in a two-hybrid assay (37). We confirmed their physical interaction with a GST pulldown assay using recombinant proteins. To map the interacting region of each protein and find a compact complex suitable for crystallization, we constructed a few fragments of each protein and tested their interaction (Figs. 1A and 4B, data not shown). These analyses show that a fragment of Faf1 spanning residues 145–220 is sufficient to bind Krr1(32–222). Moreover, Faf1(200–250) failed to bind Krr1, indicating that Faf1 residues 145–199 are the key interaction region.

We initially obtained crystals for a complex of Krr1(32–222) and Faf1(145–220), but the crystal quality was poor. After deleting the nonconserved residues 170–198 from Faf1(145–220), the crystals showed improved quality and diffracted to 2.8 Å resolution. The structure was subsequently determined by single-wavelength anomalous dispersion phasing using a selenium derivative crystal. The current model was refined to an Rwork/Rfree of 0.232/0.278 with good geometry (Table 2 and Fig. 1B and C). The asymmetric unit contains two copies of the Krr1-Faf1 complex. The two complexes show nearly identical structure and dimer interface and are superimposable with a root mean square deviation of 0.369 Å over 176 Cα pairs. Residues 38–211 of Krr1 and residues 145–163 of Faf1 were resolved in the crystal, although the other terminal residues included in crystallization fragments were invisible, likely due to structural disorder. The structure of the complex shows that the core region of Krr1 is composed of tandem KH domains, rather than one as previously thought, and the 19-residue polypeptide of Faf1 forms an α-helix that binds to the C-terminal end of the Krr1 KH domain.

Krr1 Contains Tandem KH Domains and Resembles Archaeal Dim2-like Protein—Although sequence analysis suggests that Krr1 contains only one conserved KH domain (residues 121–211, referred to as KH2), our structure shows that the region N-terminal to KH2 (residues 38–120, KH1) also adopts a KH-like fold. KH domains fall into two types of topologically distinct folds (57). The two KH domains of Krr1 belong to the type I fold, except for having an extra α-helix at the C terminus. Each domain is composed of three β-strands and four α-helices arranged in β1-α1-α2-β2-β3-α3-α4 (in the case of KH1) (Fig. 2A). The three β-strands form an antiparallel β-sheet, which packs against the four α-helices on one side. The classic KH domain is characterized by an invariant Gly-Xaa-Xaa-Gly (GXXG) sequence motif that is critical for nucleic acid interaction. The KH2 domain contains the GXXG motif that forms a short loop connecting the first two α-helices in the domain (Fig. 3A). The KH1 domain lacks the GXXG motif, and its first two α-helices are simply bent at the junction.

The two KH domains are aligned roughly in parallel such that the α3 and α4 helices of KH1 pack against the β-sheet of KH2. The extensive and mainly hydrophobic inter-domain interface literally fuses two KH domains into an integral structure. The arrangement of tandem KH domains in Krr1 is distinct from that found in other tandem KH domains, for example Nova...
KH1 and KH2 (58), but it bears a strong resemblance to that of archaeal Dim2-like proteins (PDB code 1TUA) (46, 47). The structure of Krr1 (residues 38–211) and PhDim2 (residues 33–208) can be aligned with a root-mean-square deviation of 1.28 Å over 101 Cα pairs (Fig. 3A).

KH domains generally recognize a core segment of four unpaired nucleotides in a conserved mode (59). The RNA is bound on top of an aliphatic platform formed by helix α1 and the edge of strand β2. The GXXG loop contacts the sugar-phosphate backbone of the RNA along one side. In the previously determined RNA complex structure of PhDim2, an 11-nucleotide single-stranded RNA, derived from the 3′-end of E. coli 16 S rRNA, binds across both KH domains of PhDim2 in an extended conformation (46). In the case of Krr1, the conserved KH2 domain should retain a classic RNA-binding mode. In contrast, the divergent KH1, which lacks the GXXG motif, may not bind RNA or bind in an unconventional manner.

In the aligned structures of Krr1 and PhDim2, the 5′ three nucleotides of rRNA bound to PhDim2 are incompatible with the Faf1 helix associated with Krr1 (Fig. 3A). However, the first two nucleotides do not interact with PhDim2, and the nucleotide equivalent of the third one often adopts a different conformation in the RNA complex structures of other KH domains (Fig. 3B) (59). The bound Faf1 helix might affect the conformation of the RNA at the 5′ side of the core region.

The sequence of Krr1 is extremely conserved, with 49% identity and 60% similarity between yeast and human proteins (Fig. 2A). A large portion of the Krr1 surface is covered by highly conserved residues (Fig. 3B), including the putative RNA-binding surface on KH2, the Faf1-binding channel, and the surface opposite the classic or degenerate RNA-binding sites of KH1 and KH2. This suggests that the Krr1 surface is involved in interaction with multiple partners.

**Structure of Krr1-Faf1 Complex**

Table 2

| Data collection and refinement statistics of Krr1-Faf1 structure |
|---------------------------------------------------------------|
| Values in parentheses are for the data in the highest resolution shell. r.m.s.d. means root mean square deviation. |
| Crystal form | Native | SeMet-labeled |
|----------------|--------|--------------|
| Space group | P2₁ | P2₁ |
| Cell dimensions | a, b, c (Å) | 49.1, 67.9, 81.4 | 49.1, 68.1, 81.3 |
| α, β, γ(°) | 90.0, 95.5, 90.0 | 90.0, 95.6, 90.0 |
| Wavelength (Å) | 0.9793 | 0.9793 |
| Resolution range (Å) | 50–2.8 (2.85–2.8) | 20–2.9 (2.95–2.9) |
| Unique reflections | 13,000 (559) | 11,992 (573) |
| Redundancy (I/⟨I⟩) | 3.8 (3.9) | 5.4 (3.6) |
| Completeness (%) | 99.5 (100) | 99.7 (97.8) |
| Rmerge | 0.131 (0.632) | 0.195 (0.496) |

| Structure refinement |
|---------------------|
| Resolution range (Å) | 40–2.8 (3.02–2.80) |
| No. of reflections | 12,974 (2455) |
| No. of atoms | 3161 |
| Protein | 3147 |
| Water | 14 |
| Rwork | 0.232 (0.277) |
| Rfree | 0.278 (0.350) |
| Mean B factor (Å²) | 27.8 |
| r.m.s.d. bond length (Å) | 0.003 |
| r.m.s.d. bond angles (°) | 0.644 |

**TABLE 2**

**FIGURE 2.** Multiple sequence alignments of Krr1 and Faf1. Alignment was performed for 209 Krr1 sequences (A) and 68 Faf1 sequences (B). These sequences were retrieved from the RefSeq database with blastp. Only a subset of aligned sequences are displayed. Residues that are conserved in at least 98, 80, and 60% of all aligned sequences are shaded in black, gray, and light gray, respectively. The secondary structure elements observed in the crystal structure are shown above the alignments. The residues involved in the Krr1-Faf1 interaction are indicated with circles below the alignments.
that are defective in 40 S ribosome formation at 37 °C (8, 34). Our structure provides insight into the effect of these mutations in Krr1 (Fig. 3C).

krr1–17 contains four mutations K20E, K66N, C162R, and D261A and its protein product is unstable at 37 °C. Among these mutated residues, substitution of buried Cys-162 to charged arginine may destabilize the structure. Mutation of the structurally exposed and nonconserved residue Lys-66 may be neutral. The other two residues (Lys-20 and Asp-261) are not present in the structure of Krr1 core domain and cannot be evaluated for mutational effect.

krr1–18 contains three mutations, F45L, L95S, and R207G, that are distributed in both KH domains. Mutation of buried Phe-45 to hydrophobic leucine may be tolerated structurally. The side chains of Arg-207 and Leu-95 are both exposed to the solvent, and their mutations should have a minor effect on the structure. Nevertheless, these two residues are highly conserved (Fig. 2A, 3B) and may be important for function. As the krr1–18 protein is defective in binding Kri1 (34), the mutated residues may mediate the interaction with Kri1. Leu-95 is particularly interesting because it is situated in a conserved surface patch away from the RNA-binding and Faf1-binding sites. We therefore tested whether Leu-95 is involved in binding Kri1 by a two-hybrid assay (Fig. 3D). Krr1 interacted strongly with Kri1 in the two-hybrid assay, consistent with their physical interaction previously detected by coimmunoprecipitation (34). Replacement of Leu-95 with Asp completely blocked the interaction with Kri1, whereas substitution of Asp for Phe-41, which is located near Leu-95 in structure, had no effect. This result pinpoints Leu-95 as a key site for Kri1 association.

A temperature-sensitive mutant of S. pombe Krr1, mis3-224, contains a single Gly to Glu mutation (41). The equivalent residue of yeast Krr1, Gly-175, is an invariant residue located at a
Interface are also involved in maintaining the structure of Krr1.

is extensive and flat and many hydrophobic residues at the expected to be complicated, because the binding surface of Krr1 structure and other functions of the proteins. Mutagenesis on Krr1 is involved in the Krr1-Faf1 interaction while having minimal effect on the structural helical bundle (Fig. 4

The N-terminal part of the Faf1 helix and hydrogen bond interactions and buries a solvent-accessible area of 562 Å² per subunit. The N-terminal part of the Faf1 helix, lined by Leu-155, Phe-158, Leu-159, and His-198 of Krr1, respectively. The C-terminal part of the Faf1 helix primarily contacts Krr1 via hydrophobic interactions. One face of the Faf1 helix, lined by Leu-155, Phe-158, Leu-159, and His-163, contacts with the hydrophobic surface of Krr1 comprising residues Thr-152, Leu-158, Leu-159, His-195, Ile-197, Lys-201, Ile-200, and Met-204. Most residues at the binding interface are highly conserved in their respective proteins (Figs. 2, A and B, and 3B), underscoring the importance of intermolecular association. The highly conserved hydrophobic residues Leu-149 and Leu-153, located on the exposed face of the Faf1 helix, are close to the RNA-binding surface on Krr1 KH2 and might be involved in binding RNA or other protein factors (Fig. 4A).

Functional analysis of Krr1-Faf1 interaction in yeast necessitates identification of mutants that can efficiently disrupt the Krr1-Faf1 interaction while having minimal effect on the structure and other functions of the proteins. Mutagenesis on Krr1 is expected to be complicated, because the binding surface of Krr1 is extensive and flat and many hydrophobic residues at the interface are also involved in maintaining the structure of Krr1. By contrast, Faf1 employs a short modular helix to bind Krr1, and mutation on Faf1 is more straightforward. We individually mutated three residues of Faf1, Asp-152, Leu-155, and Leu-159, and assessed the effect on Krr1 binding with a GST pulldown assay (Fig. 4B). When the protein fragments used in crystallization were assayed, the single mutations D152A and L159D and their combination reduced, but did not abolish, the interaction with Krr1, as reported previously (37). The single mutations of Faf1 failed to bind Krr1, but the D152A mutant was able to pull down Krr1. The discrepancy between the truncated and full-length proteins suggests that some residues outside the core interacting regions may be involved in intermolecular interaction.

We also assayed the mutational effect on Krr1 interaction with two-hybrid assays (Fig. 4C). The results show that Krr1 interacts strongly with Faf1, as reported previously (37). The single mutations D152A and L155D and their combination reduced, but did not abolish, the interaction with Krr1, as shown by slowed yeast growth with increasing selection stringency. Apparently, the two-hybrid assay carried out in yeast is less sensitive to these mutations than the in vitro pulldown assay. Nevertheless, deletion of the entire Krr1-binding helix or the triple mutation of D152A, L155D, and L159D (named DLL) completely abolished the Krr1-Faf1 interaction.

Krr1-Faf1 Interaction Is Essential for Cell Growth—We next asked whether the Krr1-Faf1 interaction is functionally important by using site-directed mutagenesis of Faf1 in yeast. To this end, the genomic FAF1 gene was deleted and complemented with a UR3 pRS416 plasmid expressing wild-type FAF1. This FAF1 shuffle strain was transformed with a LEU2 p415GPD
plasmid expressing HA-tagged Faf1 or mutants under the control of the constitutive GPD promoter. Growth in 5-fluoroorotic acid, which counter-selects the URA3 FAF1 plasmid, allowed the function of mutant Faf1 proteins to be assessed (Fig. 5A).

Deletion of FAF1 resulted in cell lethality, as expected for an essential gene. The D152A, L155D, and D152A/L155D mutants showed no obvious growth defect at 20, 30, or 37 °C. The triple mutant D152A/L155D/L159D and deletion of the Krr1-binding helix could not support cell growth at all three temperatures. The observed phenotypes were not due to protein instability, as all the mutant proteins were expressed at similar levels in yeast (Fig. 5B). The mutational effects on yeast growth are well correlated with the results of two-hybrid complementation and deletion strain (Fig. 5A), as reported previously (36). The growth of KM804 under non-permissive conditions can be rescued with C-terminally TAP-tagged Faf1 expressed under the constitutive GPD promoter from plasmid, but it cannot with the DLL and D145–166 mutants that are abolished in the Krr1 interaction, consistent with the complementation results conducted on the faf1 deletion strain (Fig. 5A). In addition, we tagged the genomic KRR1 gene with an HA epitope at the C terminus for protein detection; this tagging did not affect yeast growth (Fig. 6B).

The steady-state levels of rRNA processing intermediates and mature rRNAs were detected by Northern blotting using 5’-32P-labeled DNA probes (Fig. 6, C–E). Depletion of Faf1 resulted in accumulation of 35 S and 23 S pre-rRNAs and reduction of 20 S pre-rRNA (Fig. 6D, lanes 5–8). The 23 S pre-rRNA is generated from cleavage of 35 S pre-rRNA at site A3 in the absence of prior cleavage at sites A0, A1, and A2. The level of mature 18 S rRNA was also significantly reduced after 24 h of growth in glucose medium. In contrast, the level of 27 S pre-rRNA, which consists of multiple species leading to 5.8S–25 S rRNA, and mature 25 S rRNA were not affected by Faf1 depletion. These data indicate that early cleavages of 35 S pre-rRNA at sites A0, A1, and A2 were specifically inhibited in the absence of Faf1, consistent with the previous observations (36, 37).

Expression of wild-type Faf1-TAP from a p413GPD plasmid largely restored the normal processing of 18 S rRNA in cells depleted of Faf1 (Fig. 6D, lanes 1–4). However, expression of the Krr1 binding-defective mutants of Faf1-TAP failed to rescue the pre-rRNA processing (Fig. 6D, lanes 9–16). These results indicate that the Krr1-Faf1 interaction is required for early 18 S processing.

Krr1-Faf1 Interaction Is Not Required for Their Incorporation into Pre-ribosomes—The existence of interaction between Krr1 and Faf1 raises the possibility that one protein may be recruited
to pre-ribosomes by the other protein. To test this idea, the KM804/KRR1-HA strain expressing wild-type Faf1-TAP or its DLL mutant was grown in glucose to deplete the wild-type Myc-Faf1 (Fig. 7A). The cell extracts were analyzed by sucrose density gradient centrifugation, and the sedimentation behavior of Krr1-HA and Faf1-TAP was detected by Western blotting. In the presence of wild-type Faf1-TAP, Krr1 migrated broadly with a main peak at 80 S (Fig. 7B), consistent with a previous observation (8). Faf1 displayed a similar distribution pattern as Krr1, suggesting that they coexist in pre-ribosomal particles. We note that the formation of 40 S ribosome seemed not to be fully restored in this strain as the level of free 60 S subunits was high (Fig. 7B). The function of Faf1 may be affected by its C-terminal TAP tag and high expression level due to the GPD promoter.

Upon depletion of Faf1, the distribution of Krr1 in the 80 S region was not affected (Fig. 7C), indicating that Krr1 is independent of Faf1 for pre-ribosomal association. When only the DLL mutant of Faf1 was expressed, it was generally cosegregated with Krr1, yet seemed to be more accumulated at the 60 S region (Fig. 7D). The Faf1 mutant may not always assemble together with Krr1 in pre-ribosomal particles or nonspecifically associate with 60 S subunits.

The ribosomal sedimentation profiles displayed a decreased level of free 40 S ribosome and an increased level of free 60 S ribosome upon depletion of Faf1 or disruption of its interaction with Krr1. This is consistent with the defects of 18 S processing revealed above by RNA analysis.

**DISCUSSION**

Biogenesis of eukaryotic 40 S ribosome requires two functionally distinct KH proteins, Krr1 and Dim2. Krr1 functions in early 90 S pre-ribosomes, and Dim2 is present in both early 90 S and late pre-40 S pre-ribosomes. Our structure reveals that Krr1 contains a divergent KH domain in addition to the previously recognized conserved KH domain. The two KH domains of Krr1 pack into a single structural unit, which is highly similar to the structure of archaeal Dim2-like proteins. Given the significant sequence similarity with Krr1 over both divergent and conserved KH domains, Dim2 should also adopt a packed tandem KH domain structure. Dim2 has been predicted to contain a degenerate KH domain based on sequence and secondary structure analysis (60). Such an arrangement of tandem KH domains is so far unique to proteins involved in ribosome biogenesis. The presence of packed tandem KH domains in Krr1, and Dim2 by extrapolation, reinforces the notion that Krr1, Dim2, and archaeal Dim2-like proteins share a common ancestor (43). The ancestral gene likely contains two classic KH domains, similar to archaeal Dim2-like genes. After divergence of archaea and eukaryotes, the eukaryotic gene has likely lost the RNA-binding motif in KH1 and is then duplicated and
diversified to extant Krr1 and Dim2, which play different roles in ribosome biogenesis.

The conservation of RNA-binding motif in KH2 suggests that Krr1 directly binds rRNA in 90 S pre-ribosomes, although the binding target should await further characterization. Besides its role as a putative RNA-binding protein, Krr1 also serves as a protein-binding platform that interacts with a number of proteins, including the early acting 40 S synthesis factors Faf1, Kri1, and Utp14 (34, 37). In addition, Krr1 appears to form a protein module with the late-acting 40 S synthesis factors Enp1, Ltv1, Rio2, Tsr1, Dim1, and Hrr25 (61). Using structural and mutagenic approaches, we have located the binding site of Faf1 on Krr1 KH2. Our two-hybrid data also suggest that the KH1 domain of Krr1 is involved in binding Kri1, but their interaction mode should be revealed by further structural analysis. It should be noted that, in addition to tandem KH domains, Krr1 contains conserved N- and C-terminal regions (Fig. 2), which account for 45% of total residues and could contribute to protein and RNA interactions.

The KH domain has been best studied as an RNA-binding module (39). Our findings illustrate the versatility of the KH domain in protein binding. The classic KH2 domain of Krr1 uses different faces for RNA and protein binding, whereas the divergent KH1 domain of Krr1 appears to be specialized in protein recruitment. The protein-binding mode of KH2 with Faf1 is unprecedented in KH domains and distinct from that between PhDim2 and eIF2a (46). Protein binding function has been associated with other KH proteins involved in ribosome biogenesis. The divergent KH1 domain of Dim2 is implicated in Nob1 association (60), and the classic KH1 of PhDim2 was observed to interact with the translation initiation factor eIF2α (46). These examples show that the KH domain can bind protein regardless of the presence of the RNA-binding motif.

Faf1 has been shown to interact with multiple proteins besides Krr1. Faf1 demonstrates a two-hybrid interaction with Pxr1/Gno1 (37), which is also involved in ribosome biogenesis (62). Furthermore, Faf1, the 90 S component Utp11, the r-protein Rps16, and the pre-60 S factors Ebp2 and Rrp14 mutually interact with each other in two-hybrid assays (36, 63). The Faf1 sequence consists of six conserved segments of 20–40 residues in length, separated by rather variable sequences (Fig. 1). We show that the third conserved segment of Faf1, located in the middle of the protein, is responsible for binding Krr1. It is tempting to speculate that Faf1 may be a nonglobular scaffolding protein that uses separate short regions to interact with different proteins.

In contrast to the universal presence of Krr1 in eukaryotes, Faf1 exists only in fungi. Paradoxically, the Faf1-binding site is highly conserved from yeast to human Krr1, and the interaction between Krr1 and Faf1 is essential in yeast. Non-fungi eukaryotes likely have a functional equivalent of Faf1 that interacts with Krr1. However, we failed to identify any meaningful protein using the short Krr1-binding motif of Faf1 as a query in Blast.

The function of a protein in ribosome assembly is often inferred by phenotype caused by depletion of the target protein. If a protein is involved in multiple interactions, its depletion would cause complicated consequences on the assembly of pre-ribosomes (18, 19). Elucidation of the precise binding mode between Krr1 and Faf1 allowed us to study the functional role of a single protein–protein interaction, while minimizing the side effects associated with protein depletion. A subset of 90 S proteins has been shown to form subcomplexes, which are generally thought to assemble into pre-ribosomes as single entities (18, 19). However, our data show that the Krr1-Faf1 complex is
an exception to this scenario. Although Krr1 and Faf1 interact with each other, they are still able to assemble into pre-ribosomes in the absence of intermolecular interaction. Interactions with other proteins or pre-rRNA appear to be sufficient for their recruitment to pre-ribosomes, and the interaction between Krr1 and Faf1 may be secondary to their independent incorporation. Disruption of Krr1-Faf1 interaction blocked early 18S rRNA processing and yeast growth, suggesting that the Krr1-Faf1 interaction is required for the 90S pre-ribosome to adopt the functional conformation.

**Acknowledgments**—We thank Keiko Mizuta (Hiroshima University, Japan) for providing yeast strains, the staff at the Shanghai Synchrotron Radiation Facility beamline BL17U for assistance in data collection, and Jing Lu for help in yeast experiments.

**REFERENCES**

1. Phipps, K. R., Charette, J., and Baserga, S. J. (2011) The SSU processome in ribosome biogenesis—progress and prospects. *Wiley Interdiscip. Rev. RNA* 2, 1–21

2. Karstein, K. (2011) Inside the 40S ribosome assembly machinery. *Curr. Opin. Chem. Biol.* 15, 657–663

3. Kressler, D., Hurt, E., and Bassler, J. (2010) Driving ribosome assembly. *Biochim. Biophys. Acta* 1803, 673–683

4. Henras, A. K., Soudet, J., Gérus, M., Lebaron, S., Caizergues-Ferrer, M., Mougín, A., and Henry, Y. (2008) The post-transcriptional steps of eukaryotic ribosome biogenesis. *Cell. Mol. Life Sci.* 65, 2334–2359

5. Woolford, J. L., Jr., and Baserga, S. J. (2013) Ribosome biogenesis in the yeast *Saccharomyces cerevisiae*. *Genetics* 195, 643–681

6. Osheim, Y. N., French, S. L., Keck, K. M., Champion, E. A., Spasov, K., Dragon, F., Baserga, S. J., and Beyer, A. L. (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

7. Miller, O. L., Jr., and Beatty, B. R. (1969) Visualization of nucleolar genes. *Science* 164, 955–957

8. Grandi, P., Rybin, V., Bassler, J., Petfalski, E., Strauss, D., Marzioch, M., Schäfer, T., Kuster, B., Tschochner, H., Tollervey, D., Gavin, A. C., and Hurt, E. (2002) 905 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

9. Dragon, F., Gallagher, J. E., Compagnone-Post, P. A., Mitchell, B. M., Porwancher, K. A., Wehner, K. A., Wormsley, S., Settlage, R. E., Shabanowitz, J., Osheim, Y., Beyer, A. L., Hunt, D. F., and Baserga, S. J. (2002) A large nucleolar U3 ribonucleoprotein required for 18S ribosomal RNA biogenesis. *Nature* 417, 967–970

10. Schäfer, T., Strauss, D., Petfalski, E., Tollervey, D., and Hurt, E. (2003) The path from nucleolus 90S to cytoplasmic 40S pre-ribosomes. *EMBO J.* 22, 1370–1380

11. Strunk, B. S., Loucks, C. R., Su, M., Vashisth, H., Cheng, S., Schilling, J. L., Brooks, C. L., 3rd, Karbstein, K., and Skiniotis, G. (2011) Ribosome assembly factors prevent premature translation initiation by 40S assembly intermediates. *Science* 333, 1449–1453

12. Watkins, N. J., Séguault, V., Charpentier, B., Nottrott, S., Fabrizio, P., Bachi, A., Wilm, M., Rosbash, M., Branlant, C., and Lührmann, R. (2004) The 90S pre-ribosome is a multimodular structure that is assembled through a hierarchical mechanism. *Mol. Cell Biol.* 24, 2026–2036

13. Zhang, L., Lin, J., and Ye, K. (2013) Structural and functional analysis of the U3 snoRNA binding protein Rrp9. *RNA* 19, 701–711

14. Liben, J., Sun, M., and Ye, K. (2013) Structural and functional analysis of Utp23, a yeast ribosome synthesis factor with degenerate PIN domain. *RNA* 19, 1815–1824

15. Lin, J., Lu, J., Feng, Y., Sun, M., and Ye, K. (2013) An RNA-binding complex involved in ribosome biogenesis contains a protein with homology to RNA CCA-adding enzyme. *PLoS Biol.* 11, e1001669

16. Zhang, C., Lin, J., Liu, W., Chen, X., Chen, R., and Ye, K. (2014) Structure of Utp21 tandem WD domain provides insight into the organization of the UTPB complex involved in ribosome synthesis. *PLoS One* 9, e86540

17. Sasaki, T., Toh-E, A., and Kikuchi, Y. (2000) Yeast krr1p physically and
functionally interacts with a novel essential Kri1p, and both proteins are required for 40S ribosome biogenesis in the nucleolus. Mol. Cell. Biol. 20, 7971–7979.

35. Gromadka, R., and Rytka, J. (2000) The KRR1 gene encodes a protein required for 18S rRNA synthesis and 40S ribosomal subunit assembly in Saccharomyces cerevisiae. Acta Biochim. Pol. 47, 993–1005.

36. Shirai, C., Takai, T., Nariai, M., Horigome, C., and Mizuta, K. (2004) Ebp2p, the yeast homolog of Epstein-Barr virus nuclear antigen 1-binding protein 2, interacts with factors of both the 60 S and the 40 S ribosomal subunit assembly. J. Biol. Chem. 279, 25333–25338.

37. Karkusiewicz, I., Rempola, B., Gromadka, R., Grynberg, M., and Rytka, J. (2004) Functional and physical interactions of Faflp, a Saccharomyces cerevisiae nucleolar protein. Biochem. Biophys. Res. Commun. 319, 349–357.

38. Rempola, B., Karkusiewicz, I., Piekar ska, I., and Rytka, J. (2006) Fcl1p and Fcflp are novel nucleolar Saccharomyces cerevisiae proteins involved in pre-rRNA processing. Biochem. Biophys. Res. Commun. 346, 546–554.

39. Valverde, R., Edwards, L., and Regan, L. (2008) Structure and function of KH domains. FEBS J. 275, 2712–2726.

40. Chan, H. Y., Brogna, S., and O’Kane, C. J. (2001) Dribble, the Drosophila KRR1p homologue, is involved in rRNA processing. Mol. Cell. Biol. 12, 1409–1419.

41. Kondoh, H., Yuasa, T., and Yanagida, M. (2000) Mis3 with a conserved RNA binding motif is essential for ribosome biogenesis and implicated in the start of cell growth and S phase checkpoint. Genes Cells 5, 525–541.

42. Wild, T., Horvath, P., Wyler, E., Widmann, B., Badertscher, L., Zemp, I., Kozak, K., Csucs, G., Lund, E., and Kutay, U. (2010) A protein inventory of Saccharomyces cerevisiae required for 18S rRNA synthesis and 40S ribosomal subunit assembly. Mol. Cell 38, 537–547.

43. Lewis, H. A., Musunuru, K., Teplov, A., Burley, S. K., and Patel, D. J. (2000) Sequence-specific RNA binding by a Nova KH domain: implications for paraneoplastic disease and the fragile X syndrome. Cell 100, 323–332.

44. Teplov, M., Malinina, L., Darnell, J. C., Song, J., Lu, M., Abagyan, R., Musunuru, K., Teplov, A., Burley, S. K., Darnell, R. B., and Patel, D. J. (2011) Protein-RNA and protein-protein recognition by dual KH1/2 domains of the neuronal splicing factor Nova-1. Structure 19, 930–944.

45. Lovell, S. C., Davis, I. W., Arendall, W. B., 3rd, de Bakker, P. I., Word, J. M., Prisant, M. G., Richardson, J. S., and Richardson, D. C. (2003) Structure validation by Co geometry: φ, ψ and Cβ deviation. Proteins 50, 437–450.

46. Janke, C., Magiera, M. M., Rathfelder, N., Taxis, C., Reber, S., Maekawa, H., Moreno-Borchart, A., Doenges, G., Schwob, E., Schiebel, E., and Knop, M. (2004) A versatile toolbox for PCR-based tagging of yeast genes: new fluorescent proteins, more markers and promoter substitution cassettes. Yeast 21, 947–962.

47. Grishin, N. V. (2001) KH domain: one motif, two folds. Nucleic Acids Res. 29, 638–643.

48. Merl, I., Jakob, S., Ridinger, K., Hiermeier, T., Deutzmann, R., Milkeireit, P., and Tschochner, H. (2010) Analysis of ribosome biogenesis factor-modules in yeast cells depleted from pre-ribosomes. Nucleic Acids Res. 38, 3068–3080.

49. Guglielmi, B., and Werner, M. (2002) The yeast homolog of human PinX1 is involved in rRNA and small nucleolar RNA maturation, not in telomere elongation inhibition. J. Biol. Chem. 277, 35712–35719.

50. Yamada, H., Horigome, C., Okada, T., Shirai, C., and Mizuta, K. (2007) Yeast Rrp1p4p is a nucleolar protein involved in both ribosome biogenesis and cell polarity. RNA 13, 1977–1987.