Nucleic Acids Research, 2024, **52**, 431–447
https://doi.org/10.1093/nar/gkad1109
Advance access publication date: 24 November 2023
RNA and RNA-protein complexes

The Efg1–Bud22 dimer associates with the U14 snoRNP contacting the 5′ rRNA domain of an early 90S pre-ribosomal particle

Olga Beine-Golovchuk, Martina Kallas, Ruth Kunze, Sabine Griesel and Jochen Baßler*

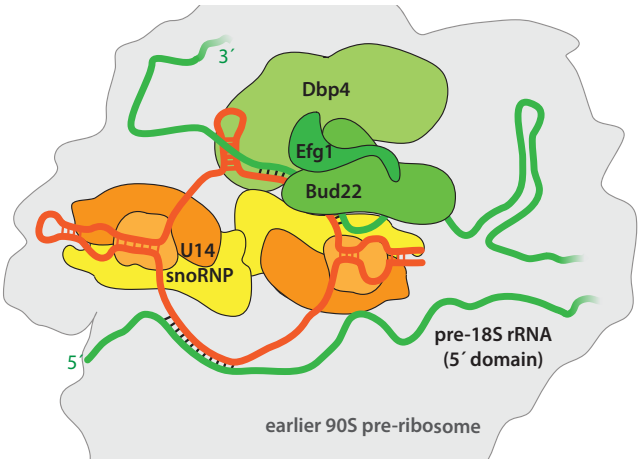
Biochemie-Zentrum der Universität Heidelberg (BZH), Im Neuenheimer Feld 328, 69120 Heidelberg, Germany

*To whom correspondence should be addressed. Tel: +49 6621 546757; Fax: +49 6621 544369; Email: jochen.bassler@bzh.uni-heidelberg.de

Abstract

The DEAD-box helicase Dbp4 plays an essential role during the early assembly of the 40S ribosome, which is only poorly understood to date. By applying the yeast two-hybrid method and biochemical approaches, we discovered that Dbp4 interacts with the Efg1–Bud22 dimer. Both factors associate with early pre-90S particles and smaller complexes, each characterized by a high presence of the U14 snoRNA. A crosslink analysis of Bud22 revealed its contact to the U14 snoRNA and the 5′ domain of the nascent 18S rRNA, close to its U14 snoRNA hybridization site. Moreover, depletion of Bud22 or Efg1 specifically affects U14 snoRNA association with pre-ribosomal complexes. Accordingly, we concluded that the role of the Efg1–Bud22 dimer is linked to the U14 snoRNA function on early 90S ribosome intermediates chaperoning the 5′ domain of the nascent 18S rRNA. The successful rRNA folding of the 5′ domain and the release of Efg1, Bud22, Dpb4, U14 snoRNA and associated snoRNP factors allows the subsequent recruitment of the Kre33-Bfr2-Enp2-Lcp5 module towards the 90S pre-ribosome.

Graphical abstract



Introduction

The assembly of ribosomes is an essential prerequisite for protein synthesis in all living cells. In eukaryotic cells, this process includes rRNA transcription, rRNA modification, rRNA folding, rRNA processing, assembly with ribosomal proteins, as well as the export to the cytoplasm. These processes are mediated by at least 200 biogenesis factors. Most of our current understanding stems from research using yeast as a model organism. Many of these biogenesis factors associate transiently with the nascent rRNA and (a subset of) ribosomal proteins to form pre-ribosomal particles, where they fulfil their structural or catalytic role. Cryo-electron microscopy (EM) analyses on purified assembly intermediates has provided detailed snapshots of ribosome biogenesis intermediates. For recent review of ribosome biogenesis in eukaryotic cells see (1–9).

Despite the functional and structural insights, a number of assembly steps are still poorly understood. One of the earliest characterized biogenesis intermediate is the 90S pre-ribosome also termed small subunit (SSU) processome. The

Received: November 21, 2022. **Revised:** October 27, 2023. **Editorial Decision:** October 30, 2023. **Accepted:** November 2, 2023

architecture of this particle contains the UTP-A and UTP-B complex, which are thought to assemble cotranscriptionally with the 5′ ETS RNA (External transcribed spacer) of the nascent rRNA. Moreover, this 90S pre-ribosome contains the U3 snoRNP, the UTP-C complex, the Mpp10-Imp3-Imp4 module, the Bms1 NTPase and its cofactor Rcl1, the Nop14-Noc4-Enp1-Emg1 subcomplex, the Rrp7-Utp22-Rrp5 module and about 15 ribosomal proteins. Finally, the 90S pre-ribosome includes the Kre33-Bfr2-Enp2-Lcp5 module and Utp20 which associate with the 5′ domain of the maturing 18S rRNA at the 90S pre-ribosome (5,10–14).

However, the presumably stepwise assembly of the 90S particle is only roughly understood (15). It is postulated that the modification of the ribosomal RNA by C/D Box snoRNP´s (methylation) and H/ACA Box snoRNP´s (pseudouridylation) occurs by binding of the snoRNP´s to the partially unfolded target rRNA sequence. It is postulated that RNA helicases mediate the release of snoRNAs, followed by the compaction of the rRNA within the 90S pre-ribosome. It remains largely elusive how rRNA modification, rRNA folding is coordinated with the assembly of the first pre-ribosomal particles.

While most snoRNA and their mediated rRNA modification are non-essential, few of them are indispensable for ribosome formation. The essential U3 snoRNA (snR17a and snR17b in *Saccharomyces cerevisiae*) does not only bind to the pre-mature 18S rRNA, but it also binds to the 5′ ETS (10,11,13,14,16). Thus, the U3 snoRNP ties together these distant rRNA sequences and keeps them in proximity. This arrangement allows the U3 snoRNA to mediate folding and stabilization of the three-dimensional structure of the nascent pre-rRNA. Subsequently, the U3 snoRNP is actively released to allow the progression of the rRNA folding. Enzymatic and structural studies revealed that the DEAD-box helicase Dhr1 promotes this detachment (17–21). A second essential snoRNA is U14 (snR128) which also belongs to the family of box C/D snoRNAs. The U14 snoRNP catalyses the methylation of C414 of the 18S rRNA which is positioned at the tip of the rRNA helix 14. The base pairing of the U14 snoRNA to the 18S rRNA helix 6 and helix14, plays a vital role by stabilizing the formation of the nascent 5′ domain of the emerging 18S rRNA. Until now, only the U3 snoRNP could be visualized in complex with its substrate by cryo-EM. For the release of the U14 snoRNP few helicases are discussed, amongst them the DEAD-box helicase Dbp4 (21,22).

Using a yeast two-hybrid approach with thermophilic proteins, combined with biochemical assays, we identified the protein interactions between Dbp4, Efg1 and Bud22. Purification of Efg1 and Bud22 associated particles contained the U3 and U14 snoRNAs, whereas the Kre33-Bfr2-Enp2-Lcp5 module that associate with the 5′ domain in later 90S particles was not yet incorporated. Further Efg1 and Bud22 were cosedimenting with lower sized particles characterized by the presence of U14 snoRNA. An RNA-crosslink analysis of Bud22 showed its binding to the U14 snoRNA and the 18S rRNA in vicinity of the U14 binding site. Moreover, deletion of Bud22 and Efg1 specifically affected the association of U14 with pre-ribosomal particles. Thus, Efg1 and Bud22 are physically and functionally linked with the U14 snoRNP during the early maturation stages of the 90S particle. We integrate our data with publish results and present a working model for early maturation of the 5′ domain.

Materials and methods

Yeast strains and growth conditions

Plasmids and yeast strains used in this study are listed in the Tables 1 and 2. Genomic tagging and gene disruptions were performed as described previously (23,24). Genomic manipulations were verified by colony PCR, Western blot analysis and/or growth phenotype. Yeast cells for tandem affinity purification, sequential proteome and RNA analyses were cultured at 30°C and harvested during logarithmic growth phase (OD_{600} of ca. 2.0–3,0). Yeast cultures for sedimentation analysis (25) were grown to OD_{600} of 0.6–0.8.

Interaction analysis for thermophilic proteins

For yeast two-hybrid (Y2H) analysis *ct*DBP4 and *ct*EFG1 alleles were N-terminally tagged with the Gal4-activation domain (AD) or the Gal4 DNA-binding domain (BD) (26). The respective plasmids pGADT7, pGBKT7 were cotransformed into the PJ69-4 MATalpha Y2H screening strain (27) and analysed for interaction by growth on SDC − Trp − Leu − His or SDC − Trp − Leu − Ade. The inhibitor of the imidazole glycerol-phosphate dehydratase 3-amino-1,2,4-triazole (3AT) was added to a final concentration of 2 mM/3 mM to avoid unspecific background growth on SDC − Trp − Leu − His plates. Cell growth was documented after 3 days of incubation at 30°C.

For biochemical reconstitution, thermophilic proteins were tagged and recombinantly expressed from plasmid in a wild-type W303 yeast strain (26). A yeast culture was grown to OD_{600} 1.0–1.5, using the appropriate selective SDC medium, harvested and resuspended in YPG medium to induce overexpression of the *Chaetomium thermophilum* genes driven by the Gal1–10 promoter. After 6 h incubation in YPG, the cells were harvested and subsequently applied to the tandem affinity-purification (TAP) procedure.

Tandem affinity purification

Two-step purification using the FTpA-tag (Flag-TEV-2x proteinA) was done according to previous protocols (28). Yeast cells of interest were mechanically milled by a cryogenic cell mill (Retsch MM400) and lysed in TRIS-based lysis buffer containing: 50 mM Tris–HCl (pH 7.5), 100 mM NaCl, 5 mM $MgCl_2$, 5% glycerol, 0.1% NP-40 and 1 mM DTT supplemented with protease inhibitor cocktail (SIGMAFAST). The lysate was cleared by centrifugation at 17 000 g for 20 min and the supernatant was incubated with IgG Sepharose 6 Fast Flow resin (GE Healthcare) during rotation for 2 h (or overnight) at 4°C. Beads were washed twice with TRIS-based washing buffer, without protease inhibitor cocktail and reduced amount of NP-40 (0.01%). Bound proteins were eluted by adding TEV protease and incubation for 1.5 h at 16°C. The obtained TEV eluate was used for binding to Flag-agarose beads (Anti-Flag M2 Affinity Gel, Sigma–Aldrich) (>1 h at 4°C). Beads bound to the protein of interest were washed with 5 ml buffer by gravity flow. Protein and associated particles were eluted with purification buffer containing Flag peptide (final concentration 300 µg/mlL). Final eluates were analysed by SDS-PAGE on 4–12% polyacrylamide gels (NuPAGE, Invitrogen) stained by colloidal Coomassie blue (Roti Blue, Roth) or analysed by northern blot analysis.

Table 1. Plasmids used in this study

Name	Number	Genotype	Reference
pGBKT7 *ctDBP4*	OY177	*GAL4* BD-*ctDBP4*, 2u, Kan, *TRP1*	(26)
pGADT7 *ctDBP4*	OY177	*GAL4* AD-*ctDBP4*, 2u, Amp, *LEU2*	(26)
pGADT7 *ctDBP4* N	6698	*GAL4* AD-*ctDBP4* N (1–402aa), 2u, Amp, *LEU2*	This study
pGADT7 *ctDBP4* N2-M1	6700	*GAL4* AD-*ctDBP4* N2-M1 (262–518aa), 2u, Amp, *LEU2*	This study
pGADT7 *ctDBP4* M-C	6697	*GAL4* AD-*ctDBP4* M-C (391–812aa), 2u, Amp, *LEU2*	This study
pGADT7 *ctEFG1*	OY102	GAL4 AD-*ctEFG1*, 2u, Amp, *LEU2*	(26)
pGBKT7 *ctEFG1*	OY102	GAL4 BD-*ctEFG1*, 2u, Kan, *TRP1*	(26)
pGBKT7 *ctEFG1* ΔN	6705	*GAL4* BD-*ctEFG1* ΔN (134–310aa), 2u, Kan, *TRP1*	This study
pGBKT7 *ctEFG1* ΔC	6706	*GAL4* BD-*ctEFG1* ΔC (1–221aa), 2u, Kan, *TRP1*	This study
pGBKT7 *ctEFG1* C	6995	*GAL4* BD-*ctEFG1* ΔC (222–310aa), 2u, Kan, *TRP1*	This study
pGBKT7 *ctBUD22*	6938	GAL4 BD-ctEFG1, 2u, Kan, *TRP1*	(26)
pGADT7 *ctBUD22*	6937	GAL4 AD-ctEFG1, 2u, Amp, *LEU2*	(26)
pMT_TRP1_(P1) P_{GAL} 3xFLAG-*ctEFG1*	OY102	2u, Amp, *TRP1*	This study
pMT_LEU2_(P1) P_{Gal} pATEV- *ctDBP4*	OY177	2u, Amp, *LEU2*	This study
pMT_TRP1_(P1) P_{Gal} 3xFLAG- *ctDBP4*	OY177	2u, Amp, *TRP1*	This study
pMT_LEU2_(P1) P_{Gal} pATEV- *ctBUD22*	OY184	2u, Amp, *LEU2*	This study
pMT_URA3_(P1) P_{Gal} HA2- *ctEFG1*	OY102	2u, Amp, *URA3*	This study
pRS315 *EFG1*	6730	*scEFG1*, *ARS/CEN*, Amp, *LEU2*	This study
pRS314 *EFG1*	6729	*scEFG1*, *ARS/CEN*, Amp, *TRP1*	This study
pRS314	9	*ARS/CEN*, Amp, *TRP1*	(55)
pRS315 *BUD22*	6732	*scBUD22*, *ARS/CEN*, Amp, *LEU2*	This study
pRS315	27	*ARS/CEN*, Amp, *LEU2*	(55)

Table 2. Yeast strains used in this study

Name	Number	Genotype	Reference
PJ69a	Y5303	*MATalpha, trp1-901, leu2-3 112, ura3-52, his3-200, gal4Δ, gal80Δ, LYS2::GAL1-HIS3, GAL2-ADE2, met2::GAL7-lacZ*	(27)
W303a	Y4392	*MATalpha, ade2-1, ura3-1, leu2-3 112, his3-11,15, trp1-1, can1-100*	(56)
FTpA-Efg1	Y5802	natNT2::ptA-TEV-Flag-P_{EFG1}, *ade2-1, ura3-1, leu2-3 112, his3-11,15, trp1-1, can1-100*	This study
Bud22-FTpA	Y6046	*BUD22*-FTpA:: natNT2, *ade2-1, ura3-1, leu2-3 112, his3-11,15, trp1-1, can1-100*	This study
Bud22-HTpA	Y5957	*BUD22*-HTpA:: *HISMX*, *ade2-1, ura3-1, leu2-3 112, his3-11,15, trp1-1, can1-100*	This study
Bud22 shuffle	Y6292	BUD22:hphNT1, *MATalpha, ade2-1, ura3-1, leu2-3 112, his3-11,15, trp1-1, can1-100*	This study
Efg1 shuffle	Y6293	*EFG1*:hphNT1, *MATalpha, ade2-1, ura3-1, leu2-3 112, his3-11,15, trp1-1, can1-100*	This study
Efg1 shuffle, Bud22Δ	Y6294	*EFG1*:hphNT1, *BUD22*::*HIS3*, *MATalpha, ade2-1, ura3-1, leu2-3 112, his3-11,15, trp1-1, can1-100*	This study
Kre33-FTpA	Y6043	*KRE33*-FTpA::*HIS3*, *ade2-1, ura3-1, leu2-3 112, his3-11,15, trp1-1, can1-100*	(12)
Cms1-FTpA	Y5887	*CMS1*-Flag-TEV-proteinA::natNT2, *MATalpha, ade2-1, ura3-1, leu2-3 112, his3-11,15, trp1-1, can1-100*	(46)
Utp10-FTpA Efg1-HA-aid	Y6295	*UTP10-FTpa::natNT2, EFG1-HA-aid::HIS3, osTIR1-myc::TRP1, ade2-1, ura3-1, leu2-3 112, his3-11,15, trp1-1, can1-100*	This study
Utp10-FTpa Bud22-HA-aid	Y6296	*UTP10-FTpa::natNT2, BUD22-HA-aid::HIS3, osTIR1-myc::TRP1, ade2-1, ura3-1, leu2-3 112, his3-11,15, trp1-1, can1-100*	This study

Sucrose gradient analysis

Sedimentation analysis of ribosomes by sucrose gradient centrifugation was performed as described previously (25). Yeast cultures were grown in 500 ml YPD medium to an OD_{600} of 0.5–0.8. Cycloheximide was added to a final concentration of approx. 100 μg/ml to block translation. After 15 min incubation with Cycloheximide, cells were harvested by centrifugation and washed in 10 ml cold buffer A (20 mM HEPES [pH 7.5], 10 mM KCl, 2.5 mM $MgCl_2$ and 1 mM EGTA). Cells were resuspended in ∼700 μl Buffer A, broken by adding glass beads and vortexing for 4 × 30 s min at 4°C (IKA Vibrax VXR). The suspension was centrifuged for 10 min at 14 000 rpm 4°C. The extracted supernatant (∼300 μl) was loaded on a 10.5 ml 10–50% sucrose gradient in buffer A and centrifuged for 16 h at 27 000 rpm using a Beckman SW40 rotor. A gradient collector (Foxy Jr from ISCO) was used to record the $UV_{254\ nm}$ absorption profile and to collect 0.5 ml fractions, which were used later for RNA extraction and Northern blot analysis or Western blot analysis. Sedimentation analysis of an FTpA-Efg1 eluate was done on 15–40% sucrose gradient based on the buffer used for the Efg1 purification. Gradient centrifugation was done using the same instruments and settings as for the yeast lysates.

RNA extraction and northern analysis

RNA obtained from different affinity-purified ribosomal particles or sucrose gradient fractions was extracted from the

Table 3. Oligonucleotides used for Northern blot analysis

Name	Sequence	Reference
anti-U3	5′-TTATGGGACTTGTT-3′	(57)
anti-U14	5′-TCACTCAGACATCCTAGG-3′	(58)
anti-snR30	5′-ATGTCTGCAGTATGGTTTTAC-3′	(59)
anti-snR83	5′-TATGAACAACAATTGTTGTAGT CGCAACTACGGTAATTGGTCC-3′	(46)

Table 4. Oligonucleotides used for CRAC analysis

Name	Sequence	Reference
L3 linker	5′-rApp/AGATCGGAAGAGCGGTT CAG/ddC/-3′	(31)
L5 Bc	5′InvddT/ACACrGrArCrGrCrUrCr UrUrCrCrGrArUrCrUrNrNrNrNrCr ArCrUrArGrC-3′	(32)
L5 Aa	5′-InvddT/ACACrGrArCrGrCrUrCr UrUrCrCrGrArUrCrUrNrNrNr Nr UrArArGrC-3′	(32)
L3 RT oligo	5′-GCTGAACCGCTCTTCCGAT-3′	(31)
P5 PCR oligo	5′-AATGATACGGCGACCACCGA GATCTACACTCTTTCCCT ACAC GACGCTCTTCCGATCT-3′	(31)
P3 PCR oligo	5′-CAAGCAGAAGACGGCATACG AGATCGGTCTCGGCATT CCTGC TGAACCGCTCTTCCGATCT-3′	(31)

eluates using phenol–chloroform followed by overnight precipitation with 70% ethanol. Alternatively, the RNA Clean & Concentrator kit from Zymo Research has been used for RNA extraction. For detection of various snoRNAs, RNA samples were separated on 6% polyacrylamide/6 M urea gels for 3 h at 400 V, followed by 40 min Sybr Green fluorescent dye staining according to the manuals (Sigma-Aldrich, S9305). For northern blot analysis, the RNA was transferred to a nylon membrane (GE Healthcare) and crosslinked at 254 nm for 2 min. The oligonucleotide probes for Northern blot analysis which have been labelled with γ-^{32}P-ATP are listed in Table 3.

Protein mass spectrometry

Indicated bands from Coomassie-stained SDS-Page gels were individually excised, trypsin digested and identified by MALDI-TOF mass spectrometry (MALDI-TOF). Semi-quantitative mass spectrometry (1D nLC–ESI-MS/MS) were done at FingerPrints Proteomics (University of Dundee, UK) and analysed by using the MaxQuant software (29). MS-data have been deposited to the ProteomeXchange Consortium via the PRIDE (30) partner repository. The dataset identifier is given in the Figure legend of Supplementary Tables S1, S2 and S5, respectively.

In vivo RNA cross-link CRAC analysis

The Protein-RNA cross-linking and cDNA analysis (CRAC) of Bud22 was performed as an *in vivo* approach as described previously with minor changes (31,32). The wild-type W303 strain was used as a control, and Bud22 was genomically tagged with His_6-TEV-ProtA at its C-terminus. Yeast cells were grown in 2l YPD media at 30°C until OD_{600} = 1.0, centrifuged and resuspended in 2 l SDC-URA media supplemented with 50 mg/l 4-thiouracil. After 4 h of further cultivation, cells were UV-irradiated for 30 min by magnetic stirrer rotation inside a crystallization dish, using UV hand lamp (365nm). Preparation of RNA and reverse transcription was done as described before (31,32) using the following reagents: RiboLock RNase inhibitor (Thermo Scientific, EO0381), TSAP (Promega, M9910), T4 PNK (NEB, M0201), T4 RNA ligase 2 truncated (NEB, M0242), T4 RNA ligase (NEB, M0204), Proteinase K (NEB, P8107), Superscript III (Invitrogen, 18080093). Used primers are listed in Table 4. The reverse transcribed and amplified cDNA from Bud22 and control sample was sequenced using the Illumina MiSeq sequencing platform and analysed using the PyCRAC software package (33) available at https://sandergranneman.bio.ed.ac.uk/pycrac-software. The used 5′ linkers contained a region of random nucleotides designed for removal of duplicate reads generated throughout PCR amplification. The obtained reads were mapped against the yeast 35S rDNA and snR128 DNA (U14) reference (see also Supplementary Table S3 and S4). The sequencing data have been deposited in NCBI's Gene Expression Omnibus (34) and are accessible through GEO Series accession number GSE214203 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE214203).

Auxin induced depletion of aid-tagged proteins

A *UTP10*-FTpA, *EFG1*-HA-aid, *osTIR1*-myc and a *UTP10*-FTpA, *BUD22*-HA-aid, *osTIR1*-myc strain was created to allow the efficient and fast depletion of *sc*Efg1 and *sc*Bud22, respectively in a Utp10-FTpA background. The addition of Auxin induces the binding of the E3 Ligase *os*Tir1 to the aid tag, which triggers its subsequent poly-ubiquitinylation. Thus, addition of Auxin to the yeast culture allows the fast and specific degradation of aid tagged proteins (35). Tagging and depletion of Efg1 and Bud22 was done as previously described (36). Auxin was added 1 h prior to the harvest of yeast cells to a final concentration of 0.5 mM. Successful depletion was monitored by Western blot analysis using an HA antibody.

Results

Identification of the Dbp4–Efg1–Bud22 module

Recently, we performed a yeast two-hybrid (Y2H) screen using a library of thermophilic ribosome biogenesis factors derived from *Chaetomium thermophilum* (*ct*) to identify previously unknown protein-protein interactions between ribosome biogenesis factors (26). In this context we identified a strong Y2H interaction between the DEAD-box helicase *ct*Dbp4 and *ct*Efg1 (Figure 1A, B), which has been linked to early steps in 40S ribosome assembly (37,38). Therefore, we were interested whether Efg1 is a regulator of the Dbp4 helicase activity. We created several truncations of *ctDBP4* and *ctEFG1* and tested them in the Y2H assay (Figure 1A, B). For *ctDBP4* we set the borders for the truncation between the N-terminal helicase core domain (^{H}N) and the C-terminal helicase core domain (^{H}C). Additional cuts were done in the Dbp4 specific C-terminus depending on the sequence conservation (see also protein alignment done with Clustal Omega (39) and Jalview (40) in Supplementary Figure S1). We found that domain 2 of the helicase core and the adjacent highly conserved region of the *ct*Dbp4 specific middle domain interacts with *ct*Efg1. We set the borders of the *ctEFG1* constructs in regions with a low amino acid conservation. We discovered that the N-terminus is dispensable for the interaction with *ct*Dbp4, whereas the C-terminus was not (Figure 1B). Moreover, a construct coding only for the C-terminus of

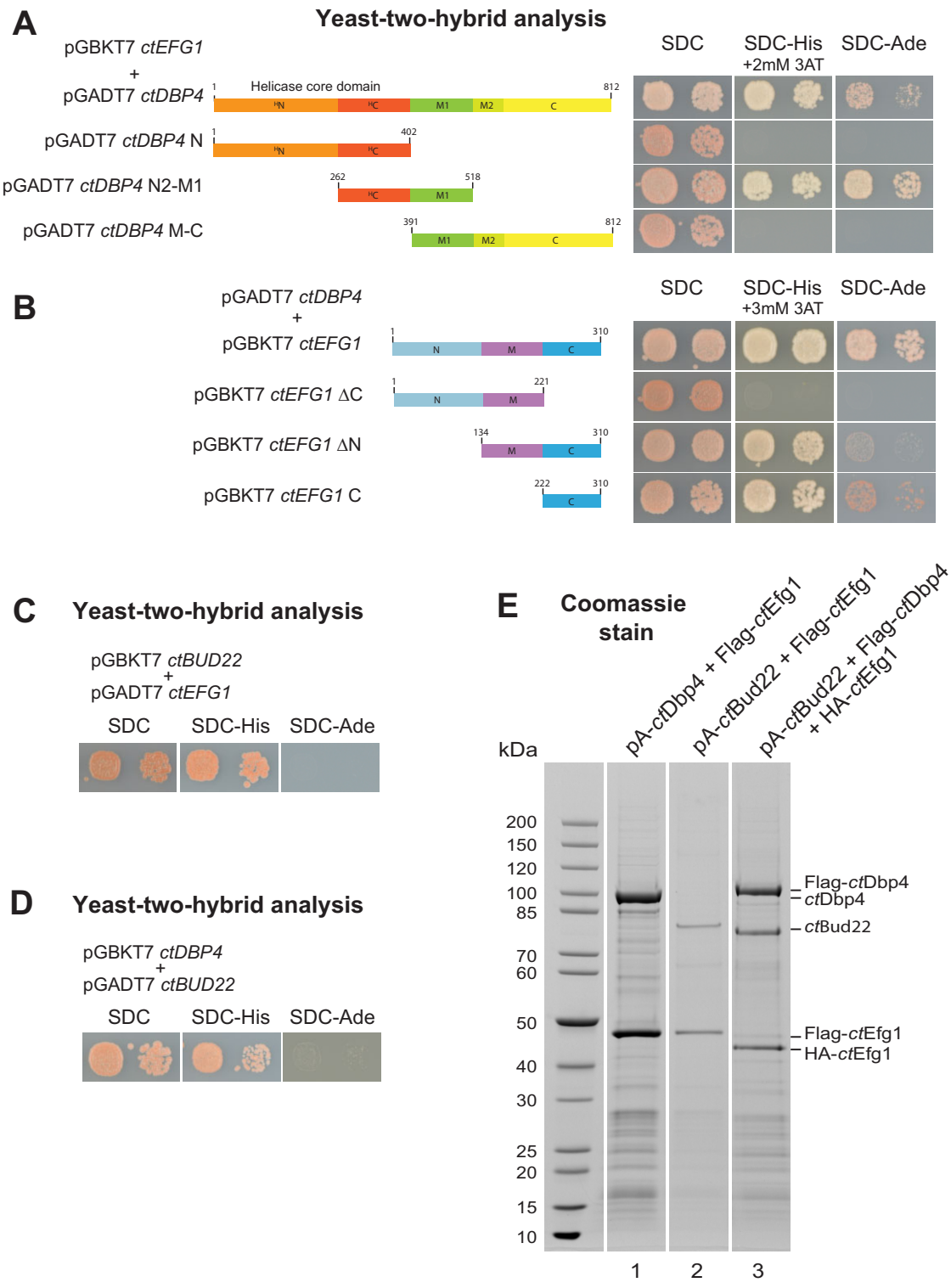


Figure 1. Interaction analysis between *ct*Efg1, *ct*Bud22 and *ct*Dbp4. (**A**) For a Yeast two-hybrid (Y2H) analysis the indicated *ctDBP4* fragments were fused to the activation domain and cotransformed with the full-length *ctEFG1* fused to the DNA binding domain into Y2H strain PJ69. Representative transformants were spotted in a 1/10 dilution series on SDC-Trp-Leu (SDC), SDC-Trp-Leu-His + 2mM 3AT (SDC-HIS) and SDC-Trp-Leu-Ade (SDC-Ade) and incubated for three days at 30°C. A schematic overview of the *ctDBP4* fragments used for the Y2H analysis is given. The Dbp4 helicase core domain consists of the N-terminal core domain (^{H}N in orange) and the C-terminal core domain (^{H}C in red). The adjacent Dbp4 specific sequence was divided into M1, M2 and C, based on the sequence conservation (See also Supplementary Figure S1). (**B**) The full-length *ctDBP4* was combined with the indicated *ctEFG1* fragments for a Y2H analysis at similar conditions as described in (A). The growth after four days of incubation at 30°C is shown. A schematic presentation of the *ctEFG1* fragments is provided. The Y2H interaction between *ctBUD22* and *ctEFG1* (**C**) and *ctBUD22* and *ctDBP4* (**D**) was analysed on SDC-Trp-Leu-His and SDC-Trp-Leu-Ade after 7 days of incubation at 30°C. (**E**) *In vitro* purification of recombinant thermophilic proteins. The proteins indicated on top of the lane were N-terminally tagged with the indicated tags and coexpressed under control of the Galactose promotor in *S. cerevisiae*. The eluates of a tandem affinity purification were analysed by 4–12% gradient SDS-PAGE and Coomassie blue staining. A protein marker is presented at the left side. Stable complex formation is demonstrated by effective coelution of the expressed proteins.

*ct*Efg1 was sufficient for the Y2H interaction with *ct*Dbp4. Thus, the required amino acid sequence of Efg1 involved in the interaction with Dbp4 did not depend on the previously crystalized central domain (37), but only on the C-terminus that was structurally not resolved. We extended this analysis by including *ct*Bud22, a poorly characterized 40S biogenesis factor (41), which was prominently copurified within Efg1 (see below). The *ct*Bud22 protein showed a weaker Y2H interaction with *ct*Efg1 (Figure 1C) and *ct*Dbp4 (Figure 1 D) as cells grew on SDC-His, but not on SDC-Ade plates. In order to test for direct biochemical interaction between *ct*Efg1and *ct*Dbp4 or *ct*Bud22, we applied a copurification assay (26). Two thermophile proteins were overexpressed in yeast, one protein carried a ptA-tag and the other protein was tagged with a Flag-tag. A subsequent two-step-affinity purification revealed, that *ct*Efg1 was capable to copurify with *ct*Dbp4 (Figure 1E, lane 1). Further, this copurification approach revealed a significant dimer formation for the ptA-c*t*Bud22 and Flag-*ct*Efg1 pair, but with a reduced protein yield in the eluate (Figure 1E lane 2). Next, we tested whether *ct*Efg1 can interact with *ct*Bud22 and *ct*Dbp4 simultaneously. Accordingly, we co-expressed ptA-*ct*Bud22, Flag-*ct*Dbp4 and HA-*ct*Efg1 in yeast and applied again the TAP protocol. This approach enabled us to isolate a trimeric thermophilic *ct*Dbp4–*ct*Efg1–*ct*Bud22 complex (Figure 1E, lane 3). However, we cannot completely rule out that an RNA molecule or protein is participating in the complex formation. In summary, our Y2H and biochemical approach revealed a biochemical interaction of Efg1 to Bud22 and Dbp4 suggesting that all three proteins cooperate together during early 40S ribosome assembly.

Efg1, Bud22 and Dbp4 associate with U14 snoRNA containing particles

Unfortunately, we were not able to uncover a direct influence of *ct*Efg1 on the *ct*Dbp4 ATPase activity *in vitro* (J.B. data not shown, see discussion). Therefore, we characterized the function of Efg1 and Bud22 *in vivo*. Because of the higher number of established genetic tools, we switched to the model organism *Saccharomyces cerevisiae* (*sc*) for further investigations. We confirmed, that the *efg1Δ* strain was sick at 30°C and below, but was not growing at 37°C (Supplementary Figure S2A) (38,42). The *bud22* deletion phenotype was less severe and showed a slow-growth phenotype at all tested temperatures (Supplementary Figure S2B). Sucrose gradient analyses revealed that deletion of *EFG1* or *BUD22* resulted in a strong reduction of 80S ribosomes and 40S subunits (see below), which is consistent with previous data (37,38,41). Consistent with the stronger growth phenotype, the *efg1* deletion strain showed a stronger reduction of 40S and 80S ribosome in a ribosome profiling analysis. Moreover, the combination of an *efg1*Δ and *bud22*Δ disruption resulted in a synthetic lethal phenotype (Figure 2A), which is consistent with an interdependent role of Efg1 and Bud22 in early 40S biogenesis. For more insight into their functional role, we analysed the sedimentation behavior of *sc*Efg1 and *sc*Bud22. This experiment revealed, that *sc*Efg1 was enriched in fractions containing 40S subunits and adjacent fraction containing smaller complexes. Moreover, *sc*Efg1 was also present in 80S ribosome containing fraction (Figure 2B). The sedimentation analysis of *sc*Bud22 showed the most prominent enrichment in lower fractions of the gradient, but was also present in fraction that contained 40S and 80S ribosomes (Figure 2B). These data indicate that Efg1 and Bud22 associate with pre-ribosomal particles. To verify this assumption, we purified *sc*Efg1 and *sc*Bud22 by the TAP method and analysed the associated proteome by mass spectrometry. Both bait proteins yielded a prominent copurification of the respective other factor (Figure 2C). Further, the helicase *sc*Dbp4 was present in both eluates visualized by Coomassie staining. The eluates showed a 90S-like protein pattern that included classic 90S biogenesis factors like members of the UTP-A and UTP-B complex, snoRNP proteins (*sc*Nop58, *sc*Nop56, *sc*Nop1, ..), *sc*Rrp5, *sc*Utp14, *sc*Mpp10. Interestingly, the *sc*Efg1 and *sc*Bud22 associated particles copurified also *sc*Dbp4, *sc*Mrd1 or *sc*Esf1, which were less frequently found in previously isolated 90S pre-ribosomal particles. In order to compare the concentration of the *sc*Efg1 and *sc*Bud22 associated proteins in more detail, we conducted a semiquantitative mass spectrometry analysis. This approach revealed that the majority of copurified proteins were present in both eluates in similar amounts (Supplementary Table S1). However, amongst the 90S associated protein factors *sc*Dbp4 was slightly enriched in the *sc*Bud22 eluate. Further *sc*Rps4, a ribosomal protein that binds early in ribosome biogenesis to the 5′ domain of the 18S rRNA was significantly enriched in the *sc*Bud22 eluate. However, we observed that the *sc*Bud22 eluate copurifies slightly higher amounts of ribosomal proteins. To determine the snoRNA content of the *sc*Efg1 and *sc*Bud22 associated particles we conducted a Northern blot analysis, which revealed the presence of the C/D-box snoRNAs U3 and U14. Whereas U3 snoRNA binds to the 5′ ETS and the H1-H2 and H27 of the 18S rRNA within the 90S particle (10,14), the U14 snoRNA is discussed to be essential for the folding of the 5′ domain of the 18S rRNA (43–45)(see also discussion). Taken together, these data suggest that the biogenesis factors Efg1 and Bud22 copurified RNP complexes included 90S pre-ribosomal particles at an early maturation stage.

Next, we tested the hypothesis that Efg1 and Bud22 are part of a precursor of the 90S pre-ribosome. Accordingly, we carefully compared a ‘late’ 90S particle, purified via the helicase and acetyltransferase *sc*Kre33 (12), with the *sc*Efg1 associated particle. A Coomassie stained SDS-PAGE gel (Figure 3A) and subsequent mass spectrometry (MS) revealed that both eluates contained a common set of protein factors like the UTP-A, UTP-B complex or C/D-Box proteins, but significantly differed for other proteins. The Kre33-associated particles included Utp20, Kre33, Bfr2, Lcp5, which were hardly detectable in the Efg1 eluate. Whereas the Efg1 purification contained specifically Bud22, Mrd1 or Esf1 (Figure 3A). A Northern blot analysis revealed that the U3 snoRNA was present at similar concentrations in both eluates, whereas the U14 snoRNA was highly prominent in the Efg1 purified eluate, but was not detectable in the Kre33 eluate (Figure 3B). In order to evaluate the differences between both particles, we applied a semi-quantitative MS analysis (Supplementary Table S2). The Label Free Quantification (LFQ) values calculated by the MaxQuant software of the purified proteins were normalized to the Utp10 value, since this UTP-A protein was present in both particles in a similar concentration (Figure 3A). These normalized values were used to calculate the factor of enrichment (Figure 3C and Supplementary Table S2). Consistent with the idea that the UTP-A and UTP-B complex associate very early with the nascent rRNA (15), the members of both modules showed no significant changes. The most drastic reduction in the Efg1 particle was observed

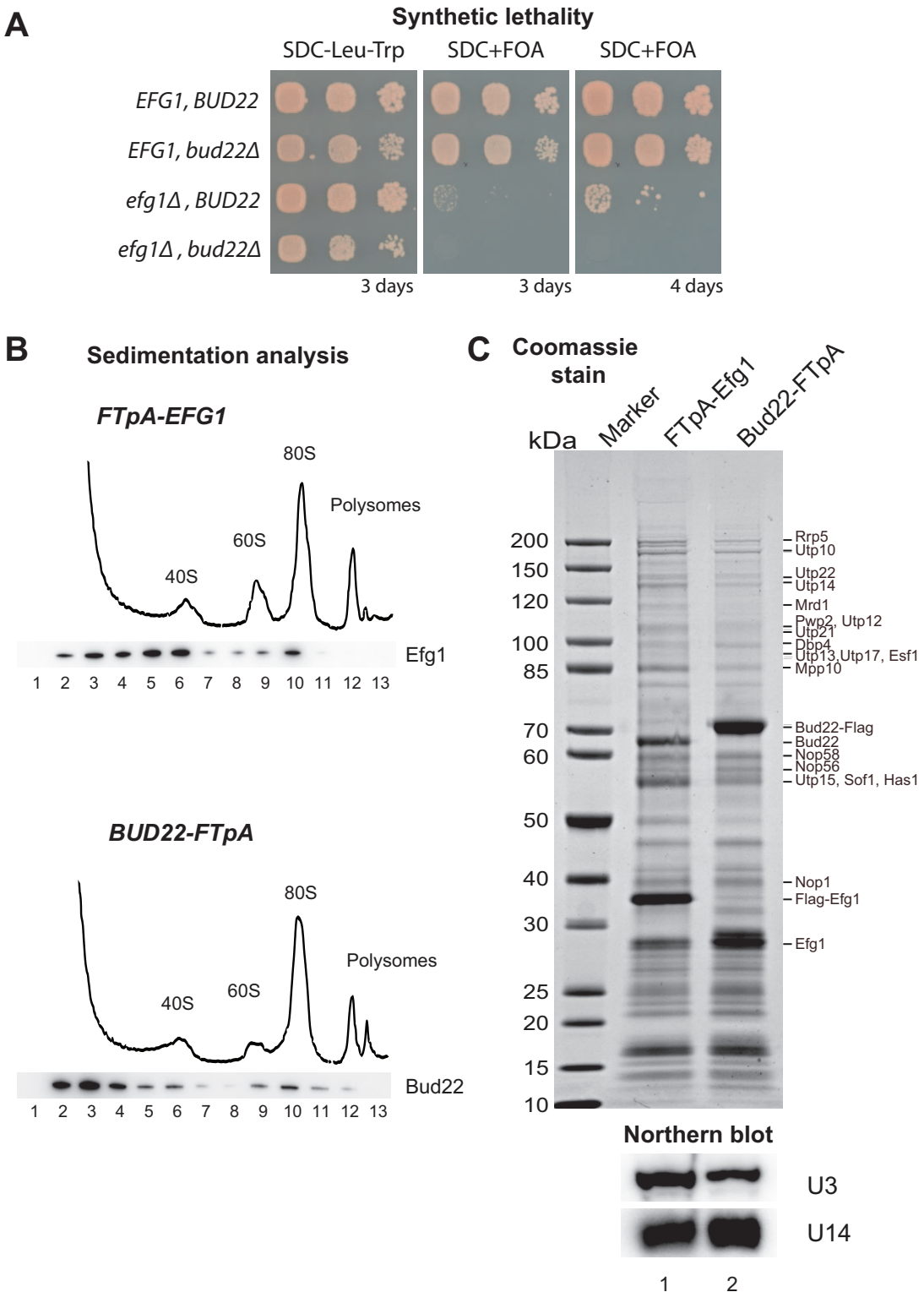


Figure 2. Functional interaction between *sc*Efg1 and *sc*Bud22. (**A**) Synthetic lethality between *sc*Efg1 and *sc*Bud22 was analysed using a double disruption strain (*bud22Δ*, *efg1Δ*) carrying a pRS316 *EFG1* (URA3) plasmid. This double deletion strain was transformed with the indicated combination of wild-type plasmids (pRS315 sc*BUD22*, pRS314 sc*EFG1*) or empty vectors. Representative transformants were grown at 30°C on SDC-Trp-Leu (3 days) and SDC + FOA plates (3, 4 days). (**B**) Sedimentation analysis of Efg1 and Bud22. A cell lysate of an FTpA-Efg1 or Bud22-FTpA expressing yeast strain was loaded on a 10–50% (w/w) sucrose gradient and ultra-centrifuged for sedimentation analysis. An absorption profile recorded at 254 nm is shown (upper panel). Derived fractions were precipitated, separated by SDS-PAGE and analysed by western blotting using an anti-ProteinA antibody (lower panel). (**C**) A tandem affinity purification was done using the yeast strains expressing FTpA-*sc*Efg1 and *sc*Bud22-FTpA, respectively. The eluates were analysed by 4–12% SDS-PAGE and Coomassie blue staining. Proteins identified by mass spectrometry are indicated at the right side, whereas the protein marker is given at the left side. A semi-quantitative mass spectrometry analysis of this purification is given in Supplementary Table S1. Aliquots of the eluates have been used for RNA extraction. Subsequent northern blot analysis was done using probes against U14 (snR128) and U3 (snR17) snoRNA and are presented in the lower panel.

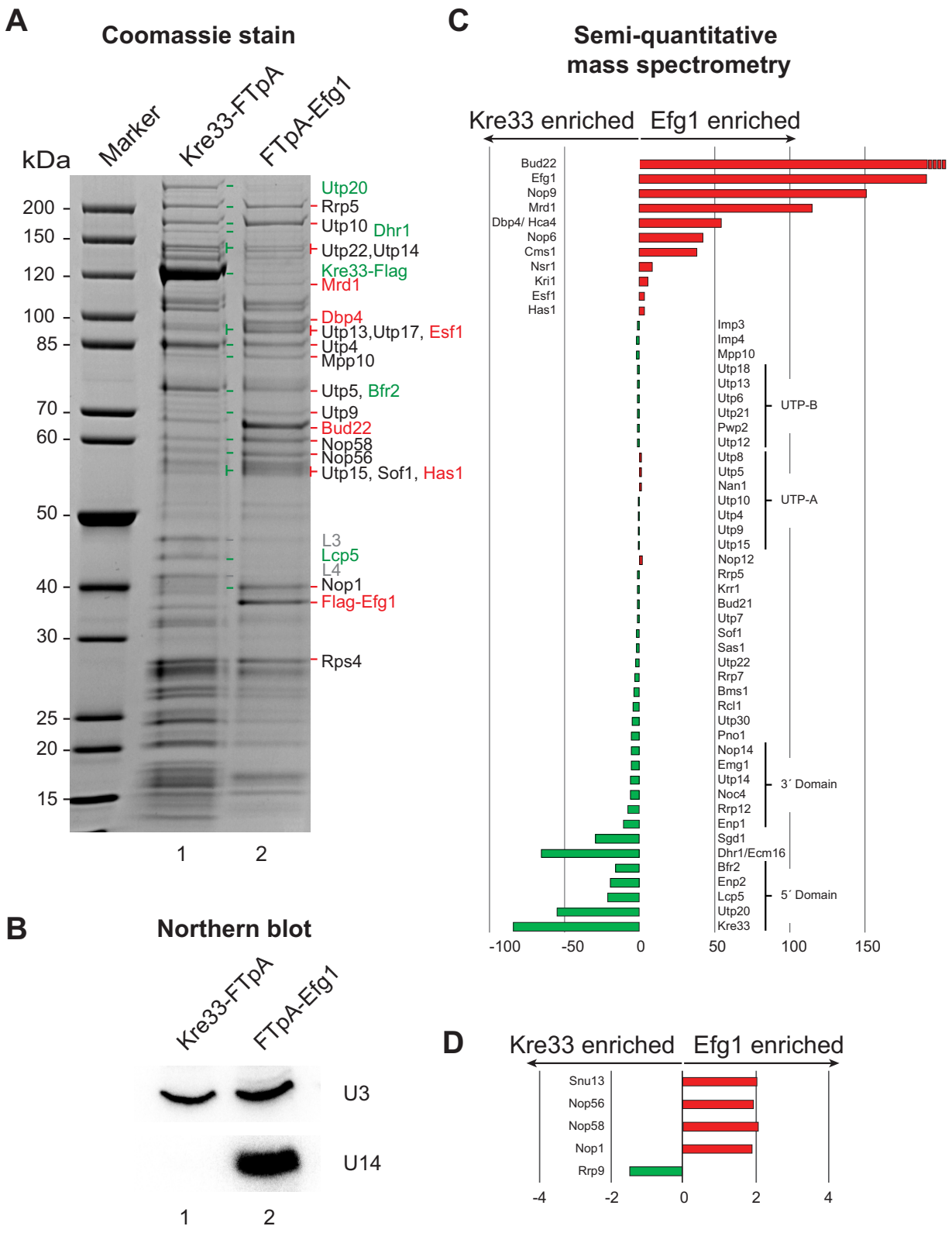


Figure 3. Comparison of a *sc*Kre33 and an *sc*Efg1 purified pre-ribosomal particle. (**A**) A genomically FTpA-tagged allele of *sc*Kre33 and *sc*Efg1, respectively, was purified using the TAP method and the resulting Flag eluate was analysed by 4–12% SDS-PAGE and Coomassie blue staining. Individual bands analysed by mass spectrometry are indicated at the right side. Common proteins determined by MALDI mass spectrometry are labelled in black, *sc*Kre33 specific proteins in green and *sc*Efg1 specific proteins in red. A protein marker is given at the left side. (**B**) The copurified RNA was analysed by northern blot analysis using probes against U3 (snr17) and U14 (snR128) snoRNA. (**C**) The protein content of both eluates was determined by semi-quantitative mass spectrometry (1D nLC-ESI-MS-MS) using the MaxQuant software (29) (see also Supplementary Table S2). The LFQ values were normalized for the biogenesis factor *sc*Utp10. The factor of enrichment (>1) in the *sc*Efg1 eluate is indicated in red (LFQ_{Efg1}/LFQ_{Kre33}). For values <1, the inverse value is displayed in green on the negative x-axis (LFQ_{Kre33}/LFQ_{Efg1}). Note that the *sc*Bud22 LFQ value was zero in the *sc*Kre33 purification. (**D**) This panel shows the factors of enrichment for the C/D box snoRNA associated proteins. Precise numbers of the semi-quantitative quantitative mass spectrometry analysis and calculations are given in the Supplementary Table S2.

for Utp20 and members of the Kre33-Enp2-Bfr2-Lcp5 module by factor of 15 or above (Figure 3C). All these biogenesis factors are bound to the 5′ domain of the nascent 18S rRNA within Kre33 associated 90S particles. Factors bound to the 3′ domain of the nascent 18S rRNA like Emg1, Enp1, Nop14 or Noc4 were only reduced by a factor of 5 to 10. Moreover, the Dhr1 helicase which releases the U3 snoRNP from late 90S particle prior to transition to the pre-40S state (17–19), seems absent in an Efg1-purified intermediate. We saw the highest enrichment for Efg1, Bud22, and Dbp4, but also a significant copurification of the very early biogenesis factors Nop9, Mrd1, Cms1 or Nop6. Remarkably, the C/D Box proteins Nop1, Nop56, Nop58 and Snu13 showed an enrichment by a factor of approximately two in the Efg1 eluate, whereas the U3 specific Rrp9 protein had a moderate reduction (Figure 3D and Supplementary Table S2). These values are nicely consistent with the observation, that the Efg1 eluate contained an additional box C/D snoRNP that harbours the U14 snoRNA. In contrast the Kre33 particle contained only the U3 snoRNP (Figure 3B). In summary, we concluded that the RNA and protein composition of the Efg1 particle is characteristic for a precursor of the 90S pre-ribosome (Kre33 particle).

In order to confirm this conclusion, we analysed the snoRNA composition of the Efg1 purified particle in more detail. For comparison, we had chosen the late factor Kre33 and the early factor Cms1, that is involved in maturation of the 3′ rRNA domain (46). Consistently, Cms1 was strongly enriched in the Efg1 purification, but reduced in the Kre33 purified 90S particle (Figure 3C). We purified Kre33, Cms1 and Efg1 associated particles and compared the snoRNA composition within the obtained eluates (Figure 4). A SYBR green staining of the RNA-polyacrylamid gel together with Northern blot analysis (Figure 4B,C) revealed that the Kre33-associated particles contained the U3 snoRNA, but only background signal of the snoRNAs U14, snR30 or snR83. In contrast, Cms1 and Efg1 copurified significant amounts of the CD box snoRNAs U3 and U14 as well as the H/ACA snoRNAs snR30 and snR83. However, the amount of coisolated snoRNA differed dramatically: Cms1 coprecipitated predominantly the snoRNA snR83, which is required for optimal maturation of the 3′ domain of early 90S intermediates (46), whereas the Efg1 eluate contained a high level of the U14 snoRNA (Figure 4B, C).

This unexpectedly large amount of copurified U14 snoRNA in the Efg1 eluate raised the possibility, whether the Efg1 eluate contains beside the 90S precursor also subcomplexes that carry the U14 snoRNA. In order to address this question, we loaded an Efg1 eluate on a sucrose gradient and determined the sedimentation of its components (Figure 5). Obtained fractions were analysed for protein content by SDS-PAGE combined with Coomassie staining (Figure 5A). Northern blot analysis was done to detect the U14 and U3 snoRNA (Figure 5B). We observed that fraction 8 contained a protein pattern characteristic for the UTP-A and UTP-B complex as well as the U3 and U14 snoRNA, which is characteristic for 90S-like particle. In the lighter fractions (Figure 5, lane 2) we detected Efg1 and Bud22 in a free form or as part of a Efg1–Bud22 heterodimer. Further, we found that fraction 4 contained high concentration of U14 snoRNA and the proteins Efg1, Bud22, Nop1 and the 5′ rRNA domain associated Rps4, amongst others. As the U3 snoRNA was absent here, we speculated that this fraction contains the U14 snoRNP that specifically associate with Efg1, Bud22 and possibly Rps4. Alternatively, this complex might represent a breakdown product containing the 5′ domain with the associated factors (see discussion). Regardless, which interpretation is correct, our data strongly suggest that Efg1 and Bud22 are located in close proximity to the U14 snoRNA on ribosomal intermediates.

Efg1 and Bud22 are associated with the 5′ domain of the nascent 18S rRNA

To confirm the position of Efg1 and Bud22 at the 5′ domain of the 18S rRNA and /or the U14 snoRNA, we conducted a crosslinking and cDNA analysis (CRAC) (31,32,47). We fed the yeast cells with Thio-uracil to increase the number of protein-RNA crosslinks and performed *in vivo* an RNA-protein crosslink with UV light (365 nm) before isolating and analysing the Efg1 and Bud22 crosslinked RNA (see materials and methods). The CRAC analysis showed that Bud22 was crosslinked to the RNA helix H9, H10 and the helices H12-H14 of the 18S rRNA which are in close proximity to the U14 binding site at the 5′ domain (Figure 6A,B and Supplementary Table S3). Beside the crosslinks to the rRNA a significant number of hits have been detected within the U14 snoRNA (Figure 6B,C and Supplementary Table S4). This cross-linked site was located at the 3′ end of the U14 snoRNA which include the 18S binding site. A CRAC analysis for Efg1 gave only preliminary data suggesting a crosslink at the 5′ rRNA domain of the 18S rRNA. Unfortunately, this data set could not be reproduced, likely because of a week affinity of Efg1 to the ribosomal RNA. Since a significant number of reproducible Bud22 hits had been found at the U14 snoRNA binding site of the 18S rRNA and U14 snoRNA, we concluded that the Efg1–Bud22 dimer binds to the hybridized U14-18S rRNA during early 90S formation. Taken together, these data are consistent with a role of Bud22 and Efg1 in the early steps of the maturation of the 5′ rRNA domain.

To find out whether Efg1 and Bud22 are functionally linked to U14 snoRNA, we investigated the sedimentation behaviour of the U14 snoRNA in *efg1* and *bud22* deletion strains. We expected that such an analysis would reveal whether U14 snoRNA binding or release from pre-ribosomal particles is affected in these knock out strains. A lysate derived from a *bud22*Δ or *efg1*Δ strain was separated on a sucrose gradient and the obtained fractions were analysed by Northern blot analysis (Figure 7). For wild-type *EFG1* and *BUD22* (Figure 7A, D), the U14 snoRNA was found to be enriched in three pools: a free pool (e.g. only associated with snoRNP proteins), a pool that comigrate with 40S subunits and a pool that comigrate with 80S ribosomes and polysomes. However, upon deletion of *BUD22* and *EFG1* (Figure 7B, E), respectively, the pool comigrating with the 40S fraction was reproducibly depleted, which might represent a smaller precursor of the 90S pre-ribosome (see discussion). Further, the amount of U14 snoRNA migrating in 80S fraction seemed to increase slightly. This effect is also seen in the relative distribution of the U14 snoRNA calculated from the intensities of the U14 Northern blot analysis (Figure 7C, F). In contrast to the altered sedimentation behaviour of the U14 snoRNA in wild-type and deletion strains, the localization of U3 snoRNA was not significantly changed in a *bud22* or *efg1* depleted situation. Hence, we concluded that Efg1 and Bud22 have a specific impact on the association of the U14 snoRNA to the pre-ribosomal intermediates.

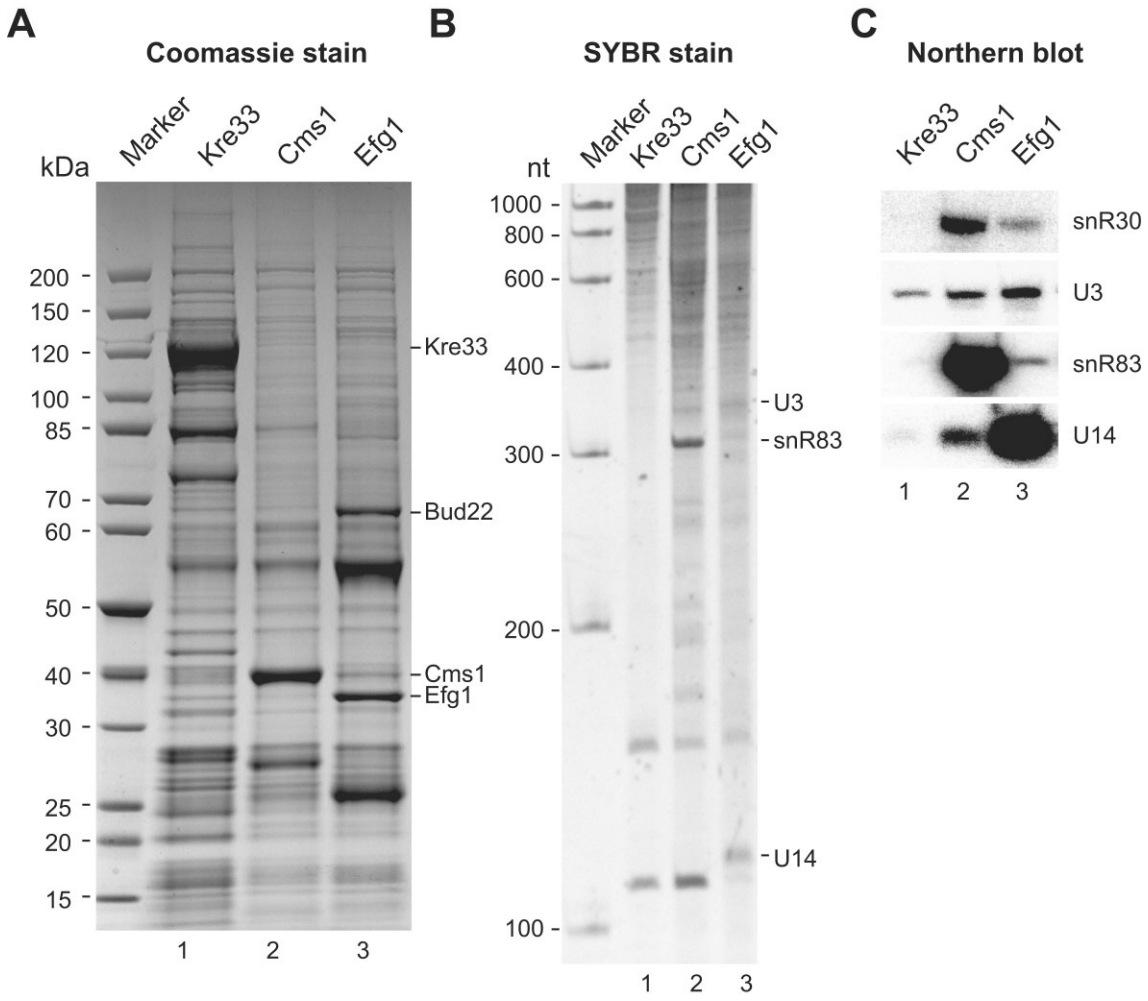


Figure 4. A pulldown of *sc*Efg1 prominently copurifies large amounts of the U14 snoRNA. (**A**) A TAP purification of *sc*Kre33, *sc*Cms1 and *sc*Efg1 was done and the eluates were analysed by 4–12% SDS-PAGE and Coomassie blue staining. A protein marker is presented at the left side, mass spectrometry of selected bands of lane 3 are indicated at the right side. (**B, C**) The RNA content of each eluate was analysed by PAGE and SYBR green staining. (C) Subsequent Northern blot analysis was done using probes against U14 (snR128), U3 (snR17), snR30 and snR83.

Next, we tested, whether the presence of Efg1 and Bud22 influences the association of the U14 snoRNA with 90S pre-ribosomal particles. We used *sc*Utp10 as bait protein, since this 90S factor was present in *sc*Efg1 and *sc*Bud22 purified particles in Coomassie stainable amounts and may not be functionally linked to Efg1 or Bud22 in ribosome assembly. We created a Utp10-FTpA, Efg1-HA-aid and a Utp10-FTpA, Bud22-HA-aid yeast strain which allows the fast and efficient depletion of the aid-tagged protein in the presence of auxin (see Materials and Methods section). The Western blot analysis confirmed the successful depletion of Efg1 and Bud22 in the cell lysate and the Utp10 eluates (Supplementary Figure S3A, C), respectively. SDS-PAGE (Supplementary Figure S3E) and semiquantitative mass spectrometry (Supplementary Table S5) indicated that eluates from the wild-type background contained a higher concentration of free UTP-A subcomplex. Moreover, depletion of Efg1 and Bud22 resulted in a slight reduction of Kre33, Cbf5 and a stronger reduction of Pol5 in the Efg1 depletion. Surprisingly, depletion of Efg1 or Bud22 does not significantly affect the 90S association of the binding partner (Supplementary Table S5) suggesting, that the recruitment of both factors is independent of the respective other factor.

The Northern blot analysis showed a significant enrichment of U14 snoRNA in Efg1 and Bud22 depleted particles (Supplementary Figure S3D), but not in the whole cell lysates (Supplementary Figure S3B). A quantification of the Northern blot confirmed an increase of the U14/U3 ration within the Utp10 purification from the depleted background. This result is consistent, with the shift of the U14 snoRNA from free and 40S fractions towards 80S fractions being observed in the sedimentation analysis in Efg1, Bud22 depleted strains (Figure 7). Thus, Efg1 and Bud22 might play a role in the release of U14 from 90S particles. However, since we were technically not able to determine the composition of the 40S sized particles carrying the U14 snoRNA, the mechanistical details of how Bud22, Efg1 and U14 snoRNA interact with each other remained unclear. Nevertheless, the here presented data support the hypothesis that Efg1, Bud22 and Dbp4 are acting together with the U14 snoRNA in the formation of the 5′ domain of the 18S rRNA on an early 90S pre-ribosome.

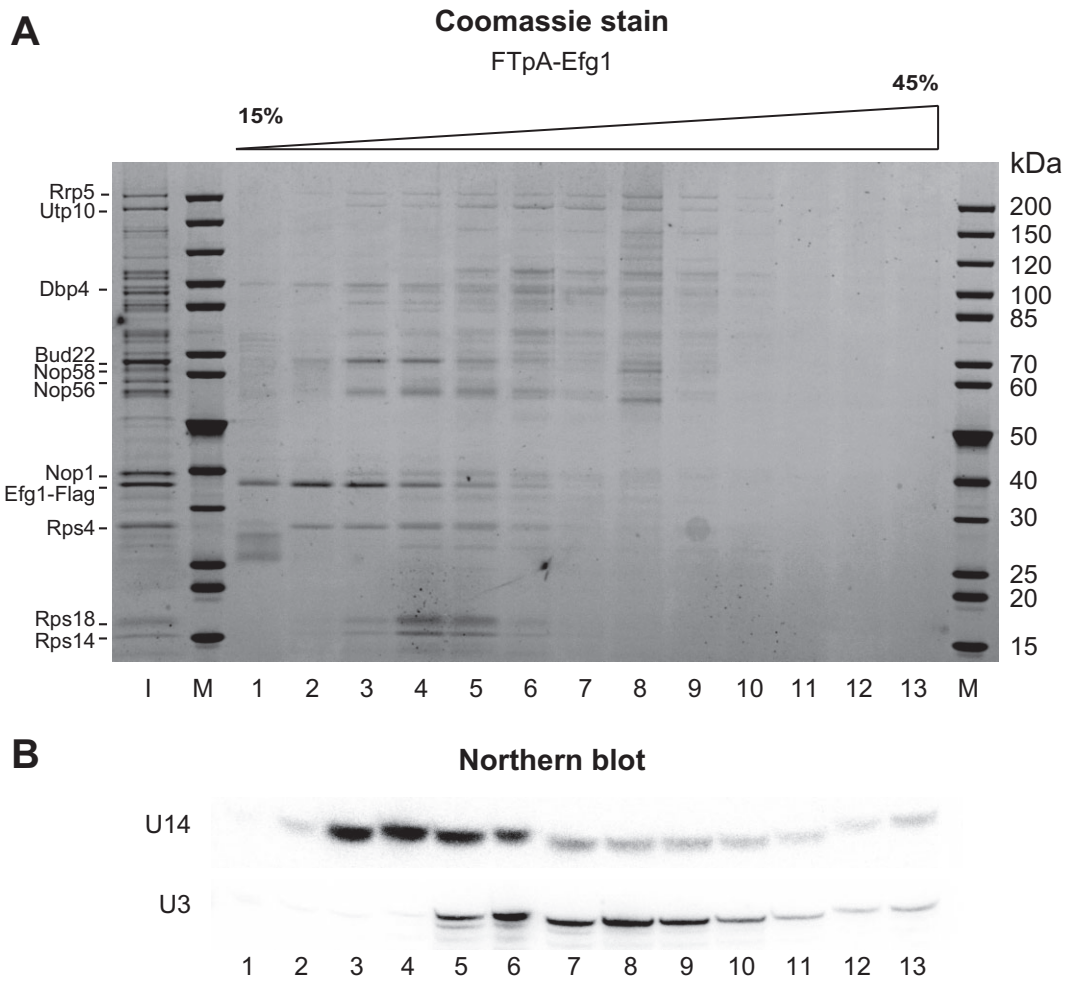


Figure 5. Efg1 and Bud22 associate with different sized particles. The eluate of an FTpA-Efg1 TAP purification was loaded on a 15–40% sucrose gradient and separated by ultracentrifugation. (**A**) Half of each fraction derived from the sedimentation analysis was precipitated by TCA and its protein content was analysed by SDS-PAGE and Coomassie staining. The input and a protein marker are shown in lane labelled with 'I' and 'M'. (**B**) The other half of the fractions was used for RNA extraction and investigated by northern blot analysis using oligos hybridizing with the U14 and U3 snoRNA. Light fractions are shown on the left side, heavy fractions are loaded on the right side of the PAGE gel.

Discussion

Using the yeast two-hybrid (Y2H) assay and a biochemically copurification method we could uncover a direct protein-protein interaction between Dbp4 and Efg1. Interestingly, Efg1 binds to the second domain of Dbp4′s helicase domain and the Dbp4 specific middle domain as seen in the Y2H assay, which prompted us to speculate that Efg1 might be a regulator of Dbp4′s helicase activity. However, by applying a malachite green based *in vitro* ATPase assay, we were not able to observe a stimulation of the *ct*Dbp4 ATPase activity, neither by adding recombinant *ct*Efg1 or RNA, nor by addition of both (Jochen Baßler, Laura Grob unpublished data). Nevertheless, the *in vivo* activation of Dbp4 might depend on its association with an early 90S particle which may carry the U14 snoRNA (see below).

Further, our biochemical data strongly suggest that Efg1 and Bud22 form a heterodimer *in vivo*. This conclusion is supported by the coexpression assay using recombinant thermophilic proteins and the *in vivo* purification of either Efg1 or Bud22 from *S. cerevisiae*. However, *in vivo* this interaction might be stabilized by RNA or a protein since we observed that the purification yield of the recombinant *ct*Efg1-*ct*Bud22 dimer is lower than for *ct*Efg1-*ct*Dbp4. In contrast, *in vivo* purifications from the yeast *S. cerevisiae* with only one tagged protein, revealed a stable dimerization between *sc*Efg1 and *sc*Bud22. Thus, potentially the *in vivo* Efg1–Bud22 binding might be stabilized by the U14 snoRNA or the ribosomal protein Rps4, which were not included within the Y2H or reconstitution assay of the thermophilic proteins. However, also different tagging might be responsible for these effects, since the *ct*Bud22 protein was tagged at its N-terminus, whereas the purifications of the yeast proteins were done with a C-terminal tagged *sc*Bud22-TAP strain.

We observed a relative strong interaction between *ct*Dbp4 and *ct*Efg1 by using a Y2H assay and a biochemical *in vitro* reconstitution approach. However, the *in vivo* purification of *sc*Efg1 from *S. cerevisiae* yielded a prominent pool of *sc*Bud22, but *sc*Dbp4 was present in lower concentrations. This discrepancy might be due to the elevated expression levels in the Y2H assay and the *in vitro* reconstitution, where strong promotors had been used. It could be also possible that the interaction of the thermophilic proteins is more stable than the mesophilic orthologues. Maybe *in vivo* Efg1 and Dbp4 might meet only during a short time window and therefore the *in vivo*

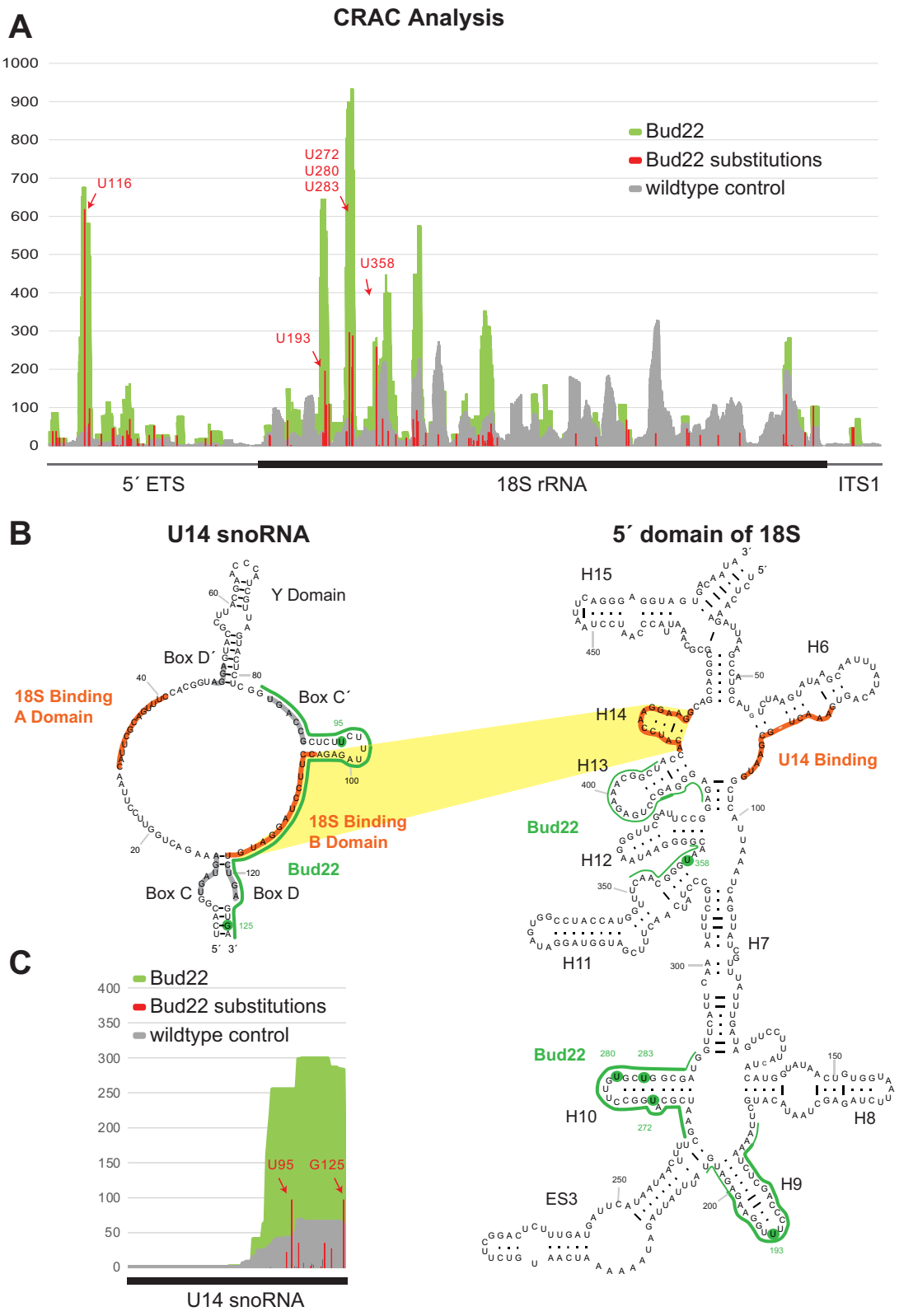


Figure 6. CRAC analysis of Bud22-HTpA. (**A**) The diagram shows the distribution of *sc*Bud22 cross-linked rRNA fragments along the 5′ ETS and 18S rRNA in green. Point mutations and deletions within the derived cDNA, indicating direct crosslinks of this nucleotide to the Bud22 protein, are shown in red, with indicated bases for prominent hits. A second dataset showing the CRAC result using a wild-type control is shown in grey. Please refer to Supplementary Table S3 for detailed analysis. (**B**) The most prominent RNA fragments crosslinked to *sc*Bud22 are indicated in green within the 2D scheme of the U14 snoRNA (left) and the 5′ domain of the 18S rRNA (right). The hybridization sites between the U14 snoRNA and the 18S rRNA are indicated in orange. Yellow background indicated the hybridization between the 18S binding B domain of the U14 snoRNA with the RNA helix H14 of the 18S rRNA. The 2D model of the U14 snoRNA is adapted from (44,60). (**C**) The diagram displays the *sc*Bud22 CRAC results matching the U14 snoRNA (green bars), the wild-type control is indicated in grey, point mutations and deletions are shown as red bars (For detailed data see also Supplementary Table S4). The number of crosslinks is shown at the same scale that in Figure 6A.

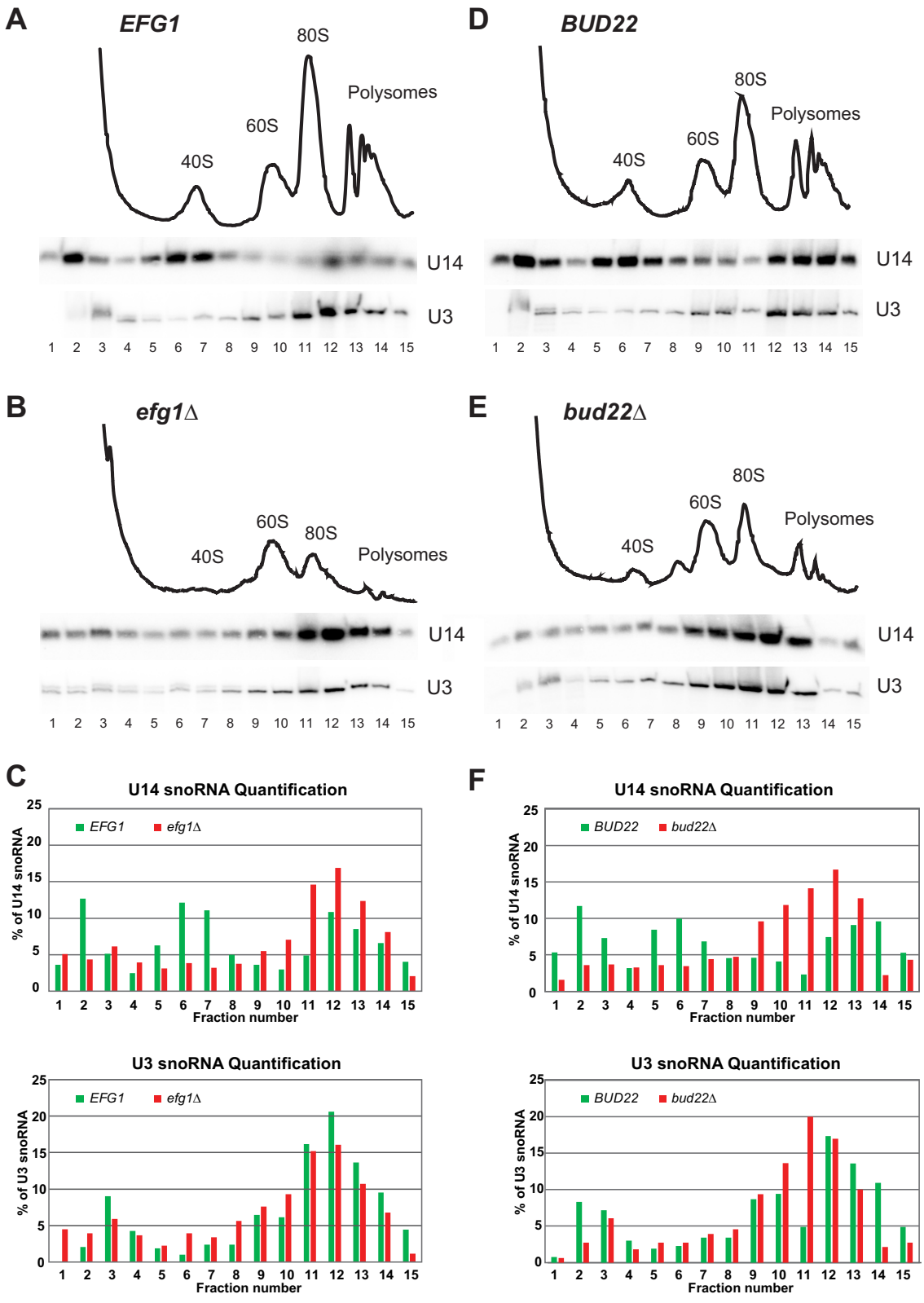


Figure 7. Deletion of *sc*Efg1 or *sc*Bud22 affect the sedimentation of the U14 but not of the U3 snoRNA. A ribosome profile analysis of the *efg1* (**B**) or *bud22* deletion (**E**) strains, with the isogenic wild-type strains (**A, D**) was performed. The sedimentation behaviour of the U14 snoRNA was compared to the migration of the U3 snoRNA (lower panel) by Northern blot analysis of the respective fractions collected from the sucrose gradient analysis. The signal intensity of the U14 snoRNA Northern blot analysis was determined and normalized to the sum of the intensity of fractions 1–15. The resulting percentage is shown in a column chart for *EFG1* and *efg1*Δ (panel C) and for *BUD22* and *bud22*Δ (panel F).

copurification is less prominent than the copurification from the overexpressed recombinant situation. Finally, an unknown regulation of the complex formation or dissociation might be the reason for the relative weak copurification of *sc*Dbp4 in the *sc*Efg1 eluate.

A previous study revealed that Efg1 localize to nucleolus and coprecipitates 35S, 23S, U3 and U14 RNA (38). In contrast, less data was available for Bud22, which was found to be involved in processing the 35S rRNA into mature 18S rRNA (41). These data are fully consistent with our findings, which additionally revealed a detailed insight into the proteome of the Efg1–Bud22 associated particles. Interestingly, an analysis of proteins associated with different truncated 18S rRNA precursors revealed that Efg1, Bud22 and Dbp4 were prominently associated with rRNA fragments comprising the 5′ domain, but were reduced in 'classic' 90S particles (48). In detail, Efg1 was suggested to bind to a minimal rRNA fragment ranging from 5′ ETS to H10, whereas the minimal rRNA fragment of Bud22 required the 5′ ETS until H13 of the nascent 18S rRNA. These results are consistent with our Bud22 crosslink data (Figure 6). Further, we discovered an elevated level of the U14 snoRNA in Bud22 and Efg1 purifications and a significant CRAC interaction of Bud22 with the U14 snoRNA. Previous EM structures of 'late' 90S pre-ribosomes revealed that the 5′ domain of the 18S rRNA is already folded and compacted, in a way that does not anymore allow binding of the U14 snoRNA. These observations support the idea that Efg1 and Bud22 interact with the 5′ domain of an early pre-90S ribosome.

Interestingly, it has been observed that a 90S purification (Noc4 bait) showed an increased concentration of U14 snoRNA when Efg1 was depleted (37), likely because of an increase of U14 snoRNA containing 90S precursors. This finding is consistent with our observation, that the concentration of the U14 snoRNA is higher within Utp10 purified particles purified from a Efg1 or Bud22 depleted background. Further, we observed that deletion of Efg1 or Bud22 abolished the association with pre-ribosomal complexes sedimenting with 40S subunits, whereas it seems that the binding of U14 snoRNA to 90S-like precursors (80S fractions) is increased. A similar shift of the U14 snoRNA from light to heavier fractions has been also observed upon depletion of the Dbp4 helicase (49). We do not know the precise nature of the 40S sized particle that contain the U14 snoRNA, but we observed that also *sc*Bud22 and especially *sc*Efg1 (Figure 2) are migrating in similar fractions in the sedimentation analysis. Our observations are consistent with previous data, as U14 snoRNA and the Dbp4 helicase were found to be enriched in 40S sized samples in former studies (49,50). A significant association with 40S sized particle has been also observed for the early biogenesis factors Krr1 (51), Nop9 (52) or Gar1 (53). All these factor act during early 40S biogenesis steps, but are absent from late pre-40S particles, which are formed by the disassembly of 90S pre-ribosomes (17,28). Therefore, we speculate that the U14 snoRNA associated particles migrating at 40S might be small precursors of the 90S pre-ribosome. However, since these very early intermediates are likely formed during ongoing transcription, this 40S sized particles could represent also particles that have been aborted from the nascent rRNA possibly during cell lysis. Unfortunately, with the current established techniques is not possible to determine the precise nature of these 40S sized particles.

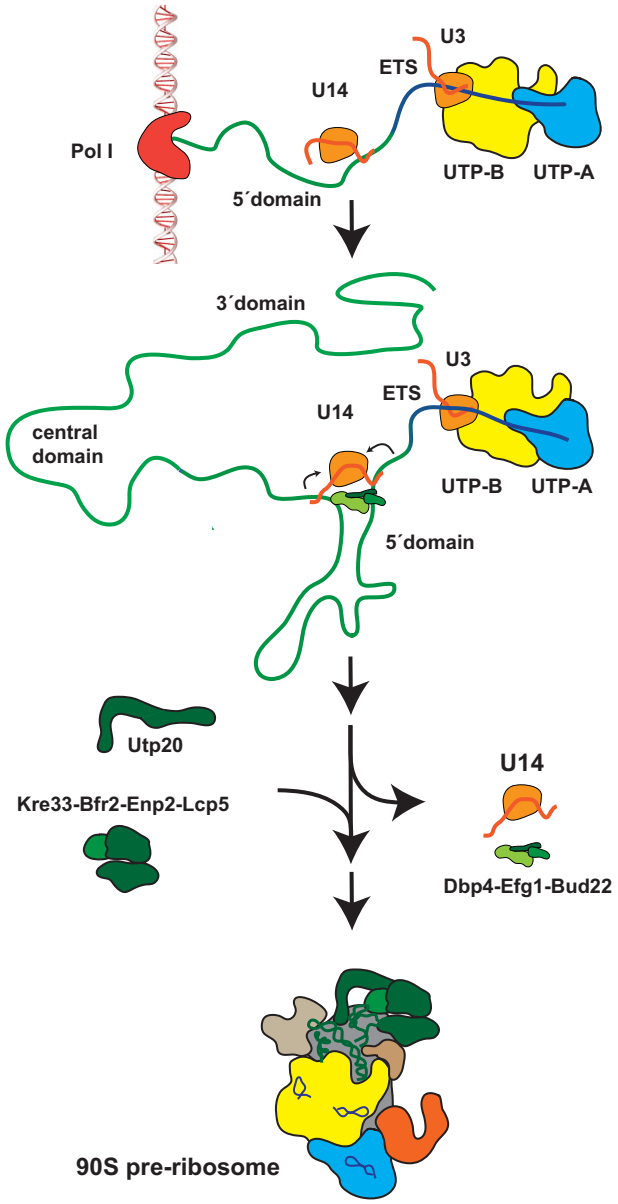


Figure 8. Working model for the role of the Efg1–Bud22 dimer and the U14 snoRNA in early steps of 90S pre-ribosome formation. For detail information see discussion.

In summary, we conclude, that the role of Efg1 and Bud22 is directly connected to the U14 snoRNP function in the maturation of the 5′ domain of the 18S rRNA. Here, our biochemical data suggest that Efg1 and Bud22 might be members an U14 specific snoRNP particle. It remains unclear, whether Efg1 and Bud22 are primarily responsible for the recruitment, the docking or the stabilization of the U14 snoRNA to the nascent rRNA. However, a role of Efg1 and Bud22 in the release of the U14 snoRNA to the 90S precursor could be also envisioned.

We observed that the ribosomal protein Rps4, which is positioned at the 5′ domain of the 18S rRNA, is specifically enriched in the Bud22 purification (Supplementary Table S1). Further, we noticed a cosedimentation of Rps4 with Bud22 and Efg1 in the light fractions when analysing the Efg1 eluate by sucrose gradient sedimentation (Figure 5). This finding

might point to a specific role of Bud22 and Efg1 for the recruitment of Rps4 to the early pre-ribosome.

Taken these previous data and our new data into account we suggest the following model (see also Figure 8). During an early stage the rRNA folding of the 5′ domain is promoted by the hybridization of U14 to the unfolded 18S rRNA helix H6 and H14. The Efg1–Bud22 dimer bind to the U14-rRNA hybrid within an early precursor of the 90S particle. Likely by the action of the Dbp4 helicase, which interact with Efg1, the release of the U14 snoRNA is initiated. This has been suggested by previous studies that revealed a connection between U14 snoRNA release and the functional Dbp4 helicase (22,49,54). We have tested in an *in vitro* release assay, whether ATP treatment of the Efg1 particle may induce a (partial) release of the U14 snoRNA. Under our conditions no U14 snoRNA release could be observed (Olga Beine-Golovchuk, unpublished data). However, it could be that the Efg1 particle was not competent for release of the U14 snoRNA because of missing factors or unfavoured RNA conformations. On the other hand, the isolated Efg1 particle might be not sufficiently homogenous or stable enough to be used unambiguously in such an *in vitro* assay. Accordingly, it has been reported that beside the copurification of 23S and 35S rRNA, U14 and U3 snoRNA a RNA breakdown product of 11S was enriched in a Efg1 purification (38). Along that line, we also observed particle instability in our preliminary EM analyses (Jochen Baßler, Olga Beine-Golovchuk, data not shown).

Our data show that Efg1–Bud22 act on the 5′ domain of early pre-90S particles, before the dissociation of the U14 snoRNA enables the 5′ domain of the 18S rRNA to adopt a near mature fold that likely is a prerequisite for the recruitment of Utp20 and the Kre33-Bfr2-Enp2-Lcp5 module. Consistent, with that model it has been reported that Efg1 depletion lead to Bud22 accumulation in 90S particles and depletion of Enp2, Bfr2, Lcp5 and Kre33 (37). Thus, we propose that the release of U14 snoRNP and likely Bud22, Efg1 which might be triggered by the Dbp4 helicase is accompanied by a major rearrangement of the 5′ rRNA domain.

Data availability

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE (30) partner repository with the dataset identifier PXD042897 (Supplementary Table S1), PXD036987 (Supplementary Table S2) and PXD042906 (Supplementary Table S5). The CRAC data discussed in this publication have been deposited in NCBI's Gene Expression Omnibus (34) and are accessible through GEO Series accession number GSE214203 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE214203).

Supplementary data

Supplementary Data are available at NAR Online.

Acknowledgements

We acknowledge the strong contribution of Maximilian Eivaskhani to this project, who died much too young. Max was responsible for some of our key findings of the work presented here and his contribution should have been rewarded with the co-first authorship.

Further, we are grateful to Prof. Dr. Cristina Paulino, Laura Grob and Dr. Patrick Heinrich for their support on this project. Jochen Baßler was funded by the German Research Foundation DFG BA 2316/2-1. M. Eivaskhani was paid by the DFG BA 2316/2-1, O. Beine-Golovchuk by the European Research Council Advanced Grant (ADG 741781 GLOWSOME) to Prof. Dr. Ed Hurt, University of Heidelberg. The authors gratefully acknowledge the data storage service SDS@hd supported by the Ministry of Science, Research and the Arts Baden-Württemberg (MWK) and the German Research Foundation (DFG) through grant INST 35/1314-1 FUGG and INST 35/1503-1 FUGG. For the publication fee we acknowledge financial support by Deutsche Forschungsgemeinschaft within the funding programme 'Open Access Publikationskosten' as well as by Heidelberg University.

Funding

Deutsche Forschungsgemeinschaft [BA 2316/2-1, INST 35/1314-1 FUGG, INST 35/1503-1 FUGG]; European Research Council [ADG 741781 GLOWSOME]. Funding for open access charge: Deutsche Forschungsgemeinschaft and University of Heidelberg.

Conflict of interest statement

None declared.

References

1. Bassler,J. and Hurt,E. (2018) Eukaryotic ribosome Assembly. *Annu. Rev. Biochem.*, **88**, 281–306.
2. Klinge,S. and Woolford,J.L. Jr (2019) Ribosome assembly coming into focus. *Nat. Rev. Mol. Cell Biol.*, **20**, 116–131.
3. Kressler,D., Hurt,E. and Bassler,J. (2017) A puzzle of life: crafting ribosomal subunits. *Trends Biochem. Sci*, **42**, 640–654.
4. Konikkat,S. and Woolford,J.L. Jr (2017) Principles of 60S ribosomal subunit assembly emerging from recent studies in yeast. *Biochem. J.*, **474**, 195–214.
5. Barandun,J., Hunziker,M. and Klinge,S. (2018) Assembly and structure of the SSU processome-a nucleolar precursor of the small ribosomal subunit. *Curr. Opin. Struct. Biol.*, **49**, 85–93.
6. Greber,B.J. (2016) Mechanistic insight into eukaryotic 60S ribosomal subunit biogenesis by cryo-electron microscopy. *RNA*, **22**, 1643–1662.
7. Nerurkar,P., Altvater,M., Gerhardy,S., Schutz,S., Fischer,U., Weirich,C. and Panse,V.G. (2015) Eukaryotic ribosome assembly and nuclear export. *Int. Rev. Cell Mol. Biol.*, **319**, 107–140.
8. Henras,A.K., Plisson-Chastang,C., O'Donohue,M.F., Chakraborty,A. and Gleizes,P.E. (2015) An overview of pre-ribosomal RNA processing in eukaryotes. *Wiley Interdiscipl. Rev. RNA*, **6**, 225–242.
9. de la Cruz,J., Karbstein,K. and Woolford,J.L. Jr (2015) Functions of ribosomal proteins in assembly of eukaryotic ribosomes in vivo. *Annu. Rev. Biochem.*, **84**, 93–129.
10. Sun,Q., Zhu,X., Qi,J., An,W., Lan,P., Tan,D., Chen,R., Wang,B., Zheng,S., Zhang,C., *et al.* (2017) Molecular architecture of the 90S small subunit pre-ribosome. *eLife*, **6**, e22086.
11. Chaker-Margot,M., Barandun,J., Hunziker,M. and Klinge,S. (2017) Architecture of the yeast small subunit processome. *Science*, **355**, eaal1880.
12. Cheng,J., Bassler,J., Fischer,P., Lau,B., Kellner,N., Kunze,R., Griesel,S., Kallas,M., Berninghausen,O., Strauss,D., *et al.* (2019) Thermophile 90S pre-ribosome structures reveal the reverse order of Co-transcriptional 18S rRNA subdomain integration. *Mol. Cell*, **75**, 1256–1269.

13. Cheng,J., Kellner,N., Berninghausen,O., Hurt,E. and Beckmann,R. (2017) 3.2-A-resolution structure of the 90S preribosome before A1 pre-rRNA cleavage. *Nat. Struct. Mol. Biol.*, **24**, 954–964.
14. Kornprobst,M., Turk,M., Kellner,N., Cheng,J., Flemming,D., Kos-Braun,I., Kos,M., Thoms,M., Berninghausen,O., Beckmann,R., *et al.* (2016) Architecture of the 90S pre-ribosome: a structural view on the birth of the eukaryotic ribosome. *Cell*, **166**, 380–393.
15. Perez-Fernandez,J., Roman,A., De Las Rivas,J., Bustelo,X.R. and Dosil,M. (2007) The 90S preribosome is a multimodular structure that is assembled through a hierarchical mechanism. *Mol. Cell. Biol.*, **27**, 5414–5429.
16. Barandun,J., Chaker-Margot,M., Hunziker,M., Molloy,K.R., Chait,B.T. and Klinge,S. (2017) The complete structure of the small-subunit processome. *Nat. Struct. Mol. Biol.*, **24**, 944–953.
17. Cheng,J., Lau,B., La Venuta,G., Ameismeier,M., Berninghausen,O., Hurt,E. and Beckmann,R. (2020) 90S pre-ribosome transformation into the primordial 40S subunit. *Science*, **369**, 1470–1476.
18. Sardana,R., Liu,X., Granneman,S., Zhu,J., Gill,M., Papoulas,O., Marcotte,E.M., Tollervey,D., Correll,C.C. and Johnson,A.W. (2015) The DEAH-box helicase Dhr1 dissociates U3 from the pre-rRNA to promote formation of the central pseudoknot. *PLoS Biol.*, **13**, e1002083.
19. Zhu,J., Liu,X., Anjos,M., Correll,C.C. and Johnson,A.W. (2016) Utp14 Recruits and activates the RNA helicase Dhr1 to undock U3 snoRNA from the preribosome. *Mol. Cell. Biol.*, **36**, 965–978.
20. Cheng,J., La Venuta,G., Lau,B., Berninghausen,O., Beckmann,R. and Hurt,E. (2022) In vitro structural maturation of an early stage pre-40S particle coupled with U3 snoRNA release and central pseudoknot formation. *Nucleic Acids Res.*, **50**, 11623–11923.
21. Mitterer,V. and Pertschy,B. (2022) RNA folding and functions of RNA helicases in ribosome biogenesis. *RNA Biology*, **19**, 781–810.
22. Kos,M. and Tollervey,D. (2005) The putative RNA helicase Dbp4p is required for release of the U14 snoRNA from preribosomes in saccharomyces cerevisiae. *Mol. Cell*, **20**, 53–64.
23. Janke,C., Magiera,M.M., Rathfelder,N., Taxis,C., Reber,S., Maekawa,H., Moreno-Borchart,A., Doenges,G., Schwob,E., Schiebel,E., *et al.* (2004) A versatile toolbox for PCR-based tagging of yeast genes: new fluorescent proteins, more markers and promoter substitution cassettes. *Yeast*, **21**, 947–962.
24. Longtine,M.S., McKenzie,A., Demarini,D.J., Shah,N.G., Wach,A., Brachat,A., Philippsen,P. and Pringle,J.R. (1998) Additional modules for versatile and economical PCR-based gene deletion and modification in Saccharomyces cerevisiae. *Yeast*, **10**, 953–961.
25. Bassler,J., Grandi,P., Gadal,O., Leßmann,T., Tollervey,D., Lechner,J. and Hurt,E.C. (2001) Identification of a 60S pre-ribosomal particle that is closely linked to nuclear export. *Mol. Cell*, **8**, 517–529.
26. Bassler,J., Ahmed,Y.L., Kallas,M., Kornprobst,M., Calvino,F.R., Gnadig,M., Thoms,M., Stier,G., Ismail,S., Kharde,S., *et al.* (2017) Interaction network of the ribosome assembly machinery from a eukaryotic thermophile. *Protein Sci.*, **26**, 327–342.
27. James,P., Halladay,J. and Craig,E.A. (1996) Genomic libraries and a host strain designed for highly efficient two-hybrid selection in yeast. *Genetics*, **144**, 1425–1436.
28. Sturm,M., Cheng,J., Bassler,J., Beckmann,R. and Hurt,E. (2017) Interdependent action of KH domain proteins Krr1 and Dim2 drive the 40S platform assembly. *Nat. Commun.*, **8**, 2213.
29. Cox,J. and Mann,M. (2008) MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.*, **26**, 1367–1372.
30. Perez-Riverol,Y., Bai,J., Bandla,C., Garcia-Seisdedos,D., Hewapathirana,S., Kamatchinathan,S., Kundu,D.J., Prakash,A., Frericks-Zipper,A., Eisenacher,M., *et al.* (2022) The PRIDE database resources in 2022: a hub for mass spectrometry-based proteomics evidences. *Nucleic Acids Res.*, **50**, D543–D552.
31. Granneman,S., Kudla,G., Petfalski,E. and Tollervey,D. (2009) Identification of protein binding sites on U3 snoRNA and pre-rRNA by UV cross-linking and high-throughput analysis of cDNAs. *Proc. Natl. Acad. Sci. U.S.A.*, **106**, 9613–9618.
32. Thoms,M., Thomson,E., Bassler,J., Gnadig,M., Griesel,S. and Hurt,E. (2015) The exosome is recruited to RNA substrates through specific adaptor proteins. *Cell*, **162**, 1029–1038.
33. Webb,S., Hector,R.D., Kudla,G. and Granneman,S. (2014) PAR-CLIP data indicate that Nrd1-Nab3-dependent transcription termination regulates expression of hundreds of protein coding genes in yeast. *Genome Biol.*, **15**, R8.
34. Edgar,R., Domrachev,M. and Lash,A.E. (2002) Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res.*, **30**, 207–210.
35. Nishimura,K., Fukagawa,T., Takisawa,H., Kakimoto,T. and Kanemaki,M. (2009) An auxin-based degron system for the rapid depletion of proteins in nonplant cells. *Nat. Methods*, **6**, 917–922.
36. Matsuo,Y., Granneman,S., Thoms,M., Manikas,R.G., Tollervey,D. and Hurt,E. (2014) Coupled GTPase and remodelling ATPase activities form a checkpoint for ribosome export. *Nature*, **505**, 112–116.
37. Shu,S. and Ye,K. (2018) Structural and functional analysis of ribosome assembly factor Efg1. *Nucleic Acids Res.*, **46**, 2096–2106.
38. Choque,E., Schneider,C., Gadal,O. and Dez,C. (2018) Turnover of aberrant pre-40S pre-ribosomal particles is initiated by a novel endonucleolytic decay pathway. *Nucleic Acids Res.*, **46**, 4699–4714.
39. Madeira,F., Park,Y.M., Lee,J., Buso,N., Gur,T., Madhusoodanan,N., Basutkar,P., Tivey,A.R.N., Potter,S.C., Finn,R.D., *et al.* (2019) The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Res.*, **47**, W636–W641.
40. Waterhouse,A.M., Procter,J.B., Martin,D.M., Clamp,M. and Barton,G.J. (2009) Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics*, **25**, 1189–1191.
41. Dakshinamurthy,A., Nyswaner,K.M., Farabaugh,P.J. and Garfinkel,D.J. (2010) BUD22 affects Ty1 retrotransposition and ribosome biogenesis in Saccharomyces cerevisiae. *Genetics*, **185**, 1193–1205.
42. Kastenmayer,J.P., Ni,L., Chu,A., Kitchen,L.E., Au,W.C., Yang,H., Carter,C.D., Wheeler,D., Davis,R.W., Boeke,J.D., *et al.* (2006) Functional genomics of genes with small open reading frames (sORFs) in S. cerevisiae. *Genome Res.*, **16**, 365–373.
43. Liang,W.Q. and Fournier,M.J. (1995) U14 base-pairs with 18S rRNA: a novel snoRNA interaction required for rRNA processing. *Genes Dev.*, **9**, 2433–2443.
44. Chagot,M.E., Quinternet,M., Rothe,B., Charpentier,B., Coutant,J., Manival,X. and Lebars,I. (2019) The yeast C/D box snoRNA U14 adopts a "weak" K-turn like conformation recognized by the Snu13 core protein in solution. *Biochimie*, **164**, 70–82.
45. Samarsky,D.A., Schneider,G.S. and Fournier,M.J. (1996) An essential domain in Saccharomyces cerevisiae U14 snoRNA is absent in vertebrates, but conserved in other yeasts. *Nucleic Acids Res.*, **24**, 2059–2066.
46. Lau,B., Beine-Golovchuk,O., Kornprobst,M., Cheng,J., Kressler,D., Jady,B., Kiss,T., Beckmann,R. and Hurt,E. (2022) Cms1 coordinates stepwise local 90S pre-ribosome assembly with timely snR83 release. *Cell Rep.*, **41**, 111684.
47. Bohnsack,M.T., Martin,R., Granneman,S., Ruprecht,M., Schleiff,E. and Tollervey,D. (2009) Prp43 bound at different sites on the pre-rRNA performs distinct functions in ribosome synthesis. *Mol. Cell*, **36**, 583–592.
48. Zhang,L., Wu,C., Cai,G., Chen,S. and Ye,K. (2016) Stepwise and dynamic assembly of the earliest precursors of small ribosomal subunits in yeast. *Genes Dev.*, **30**, 718–732.
49. Soltanieh,S., Osheim,Y.N., Spasov,K., Trahan,C., Beyer,A.L. and Dragon,F. (2015) DEAD-box RNA helicase Dbp4 is required for small-subunit processome formation and function. *Mol. Cell. Biol.*, **35**, 816–830.
50. Soltanieh,S., Lapensee,M. and Dragon,F. (2014) Nucleolar proteins Bfr2 and Enp2 interact with DEAD-box RNA helicase Dbp4 in two different complexes. *Nucleic Acids Res.*, **42**, 3194–3206.

51. Grandi,P., Rybin,V., Bassler,J., Petfalski,E., Strauss,D., Marzioch,M., Schäfer,T., Kuster,B., Tschochner,H., Tollervey,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.
52. Thomson,E., Rappsilber,J. and Tollervey,D. (2007) Nop9 is an RNA binding protein present in pre-40S ribosomes and required for 18S rRNA synthesis in yeast. *RNA*, **13**, 2165–2174.
53. Lemay,V., Hossain,A., Osheim,Y.N., Beyer,A.L. and Dragon,F. (2011) Identification of novel proteins associated with yeast snR30 small nucleolar RNA. *Nucleic Acids Res.*, **39**, 9659–9670.
54. Liang,W.Q., Clark,J.A. and Fournier,M.J. (1997) The rRNA-processing function of the yeast U14 small nucleolar RNA can be rescued by a conserved RNA helicase-like protein. *Mol. Cell. Biol.*, **17**, 4124–4132.
55. Sikorski,R.S. and Hieter,R. (1989) A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in saccharomyces cerevisiae. *Genetics*, **122**, 19–27.
56. Thomas,B.J. and Rothstein,R. (1989) Elevated recombination rates in transcriptionally active DNA. *Cell*, **56**, 619–630.
57. Sharma,K. and Tollervey,D. (1999) Base pairing between U3 small nucleolar RNA and the 5' end of 18S rRNA is required for pre-rRNA processing. *Mol. Cell. Biol.*, **19**, 6012–6019.
58. Morrissey,J.P. and Tollervey,D. (1997) U14 small nucleolar RNA makes multiple contacts with the pre-ribosomal RNA. *Chromosoma*, **105**, 515–522.
59. Fath,S., Milkereit,P., Podtelejnikov,A.V., Bischler,N., Schultz,P., Bier,M., Mann,M. and Tschochner,H. (2000) Association of yeast RNA polymerase I with a nucleolar substructure active in rRNA synthesis and processing. *J. Cell Biol.*, **149**, 575–590.
60. Samarsky,D.A., Fournier,M.J., Singer,R.H. and Bertrand,E. (1998) The snoRNA box C/D motif directs nucleolar targeting and also couples snoRNA synthesis and localization. *EMBO J.*, **17**, 3747–3757.

Received: November 21, 2022. **Revised:** October 27, 2023. **Editorial Decision:** October 30, 2023. **Accepted:** November 2, 2023