Structure
Article

Comprehensive Molecular Structure of the Eukaryotic Ribosome

Derek J. Taylor,[1,5] Batsal Devkota,[2,6] Andrew D. Huang,[2] Maya Topf,[3,7] Eswar Narayanan,[3] Andrej Sali,[3] Stephen C. Harvey,[2] and Joachim Frank[4,*]
[1]Wadsworth Center, Empire State Plaza, Albany, NY 12201-0509, USA
[2]School of Biology and Parker H. Petit Institute of Bioengineering and Biosciences, Georgia Institute of Technology, Atlanta, GA 30332-0230, USA
[3]Department of Bioengineering and Therapeutical Sciences, Department of Pharmaceutical Chemistry, and California Institute of Quantitative Biosciences, Mission Bay Byers Hall, University of California at San Francisco, San Francisco, CA 94158, USA
[4]Howard Hughes Medical Institute, Department of Biochemistry and Molecular Biology and Department of Biological Sciences, Columbia University, New York, NY 10032, USA
[5]Present address: Department of Pharmacology, Case Western Reserve University, Cleveland, OH 44106-4965, USA
[6]Present address: Research Collaboratory for Structural Biology, Rutgers, State University of New Jersey, Department of Chemistry and Chemical Biology, Piscataway, NJ 08854-8087, USA
[7]Present address: Institute of Structural and Molecular Biology, School of Crystallography, Birkbeck College London, London WC1E 7HX, UK
*Correspondence: jf2192@columbia.edu
DOI 10.1016/j.str.2009.09.015

SUMMARY

Despite the emergence of a large number of X-ray crystallographic models of the bacterial 70S ribosome over the past decade, an accurate atomic model of the eukaryotic 80S ribosome is still not available. Eukaryotic ribosomes possess more ribosomal proteins and ribosomal RNA than do bacterial ribosomes, which are implicated in extraribosomal functions in the eukaryotic cells. By combining cryo-EM with RNA and protein homology modeling, we obtained an atomic model of the yeast 80S ribosome complete with all ribosomal RNA expansion segments and all ribosomal proteins for which a structural homolog can be identified. Mutation or deletion of 80S ribosomal proteins can abrogate maturation of the ribosome, leading to several human diseases. We have localized one such protein unique to eukaryotes, rpS19e, whose mutations are associated with Diamond-Blackfan anemia in humans. Additionally, we characterize crucial interactions between the dynamic stalk base of the ribosome with eukaryotic elongation factor 2.

INTRODUCTION

The ribosome is a massive ribonucleoprotein particle composed of a small and a large subunit (40S and 60S in eukaryotes; 30S and 50S in eubacteria). The eukaryotic ribosome is a complex assembly of four ribosomal RNAs (rRNA) with about 80 ribosomal proteins (rps), and its assembly requires more than 150 nonribosomal factors (Ferreira-Cerca et al., 2005). Although a preeminent role of the ribosome is to translate messenger RNA into polyeptides, it also acts as a platform for several nonribosomal proteins involved in fundamental biological processes, including docking of the ribosome to cellular organelles and recruiting kinases engaged in various signaling pathways. In rapidly growing yeast cells, the majority of transcription is devoted to the production of ribosomal RNA, and about 50% of RNA polymerase II transcription is committed to production of ribosomal proteins (Warner, 1999), highlighting the overall importance of ribosome function within the cell.

Because of the vast amount of available genetic information and the simplicity of its system, compared with those of other eukaryotic cells, the yeast *Sacchromyces cerevisiae* has been studied extensively as a model system for eukaryotes. Most of our knowledge about eukaryotic ribosome biogenesis derives from genetic studies in yeast. All the ribosomal proteins in yeast have homologs in ribosomes from higher-order organisms, making the yeast ribosome an excellent model for characterizing eukaryotic ribosome structure and function, as well as protein synthesis and ribosome biogenesis. The small ribosomal subunit in yeast is composed of an 18S rRNA and 32 rps, whereas the 60S subunit is composed of 3 rRNAs and 46 rps. The rRNA "core" of the 80S ribosome is homologous to that of the bacterial 70S ribosome, with a few modifications. The 80S ribosome contains a novel 5.8S rRNA containing 158 nt; it is homologous to the 5′ end of 23S rRNA in the 70S ribosome. The 25S rRNA (3392 nt) in the 80S is homologous with the remaining 3′ sequence of the 23S bacterial rRNA. Although it is homologous, the 5.8S/25S rRNA contains approximately 20% more bases than does the 23S rRNA in 70S ribosomes. The last piece of rRNA in the 60S subunit is a 5S rRNA (121 nt in yeast), which is conserved among ribosomes from all kingdoms. The eukaryotic 40S subunit consists of a single piece of 18S rRNA (1798 nt) that is about 15% longer than, but homologous to, 16S rRNA in bacteria.

The core of the ribosome—where aminoacylated-tRNA is decoded, peptide bonds are formed, and tRNAs are translocated from defined binding sites of the ribosome to the next—is highly conserved in the ribosomes of all species of life, be it plant, animal, or bacteria (Gutell et al., 1985). This so-called catalytic core is conserved across species, but eukaryotic ribosomal RNA contains more than 50 additional nucleotide sequences called expansion segments inserted at specific positions of the

conserved rRNA core (Gerbi, 1996). The 18S rRNA contains 12 expansions segments (ES) over its 16S rRNA bacterial counterpart, and there are 41 ES located throughout the rRNA of the large subunit (Gerbi, 1996). Complete removal of ES27 from *Tetrahymena thermophila* rRNA in the large subunit deleteriously affects processing and stability of the rRNA (Sweeney et al., 1994). Similarly, inserting short sequences into ES3 of 18S rRNA in yeast interferes with ribosome assembly (Musters et al., 1990). For the most part, however, the role of expansion segments has not been well characterized.

The ribosomal proteins appear to play a supporting role in protein synthesis but are also required for accurate cleavage and processing of rRNA during ribosome assembly, nuclear export, and maturation of the ribosome (Wool, 1996). Most of the 80S rps are absolutely required for viability. In fact, it has been reported that at least 28 of the 40S rps are necessary for cell growth (Ferreira-Cerca et al., 2005). A similar percentage of rps in the 60S subunit required for viability can be assumed as well. After being translated in the cytoplasm, ribosomal proteins are imported into the nucleus and then the nucleolus, where they assist in rRNA folding and maturation, as well as subunit biogenesis and nuclear export, whereupon the preribosome is further processed in the cytoplasm (Leger-Silvestre et al., 2004). All this suggests that the rps in eukaryotes play a critical role in extra-ribosomal functions, ribosome biosynthesis, and eukaryotic trafficking.

The eubacterial 70S ribosome and the individual small subunit, as well as the archaeal large subunit, have been solved to atomic resolution (Blaha et al., 2009 and references therein). Although our understanding of ribosome structure and function has been greatly enhanced by the three-dimensional structures, detailed knowledge is still largely limited to bacterial ribosomes. Several new and expanded molecules reside in the eukaryotic ribosome, making it significantly more complex. In the small subunit, only 15 of the 32 ribosomal proteins, based on orthologs of the available crystal structures from eubacteria, have been described in the three-dimensional context of the eukaryotic ribosome from yeast (Spahn et al., 2001; Spahn et al., 2004). The large subunit consists of 46 and 33 proteins in yeast and *Escherichia coli*, respectively, with 18 of those proteins sharing significant sequence homology in the two systems. The sequence homology of these proteins is extended to include those in the mammalian ribosome as well (Chandramouli et al., 2008). Modeling of eukaryotic ribosomal RNA (rRNA) has also been limited to docking of available crystal structures from eubacterial and archaeal sources, omitting rRNA expansion segments unique to the 80S ribosome (Spahn et al., 2001, 2004). Recent progress has been made to dock rRNA sequences into the cryo-EM map of a dog ribosome, but the model is still missing approximately 50% of the rRNA sequence attributed to expansion segments (Chandramouli et al., 2008).

To our knowledge, we present here the most complete structure of the eukaryotic ribosome to date. Our study utilizes the 8.9 Å cryo-EM reconstruction of the 80S ribosome from the thermophilic fungus *Thermomyces lanuginosus* (Taylor et al., 2007) as a guide for building homology models of the ribosomal proteins and rRNA sequences that fit into the cryo-EM map. Because the *T. lanuginosus* ribosome shares over 85% sequence identity with that of *S. cerevisiae* (Nilsson et al., 2007), we have used the sequence of rRNA and ribosomal proteins from *S. cerevisiae* for modeling and docking.

The intrinsic dynamic motions of ribosomal substructures, such as the L1 stalk and the stalk-base of the ribosome, have made these regions particularly difficult to characterize by using X-ray crystallographic methods. The structures of atomic models docked into our cryo-EM map allow interpretation of these regions that have not been well characterized to date. Such motions include those of the L1 stalk, the stalk base, and the ratcheting motion of the small subunit relative to the large subunit. Finally, the 80S ribosome presented here is complexed with the eukaryotic elongation factor 2 (eEF2), the homolog of EF-G in bacteria, which catalyzes translocation of tRNAs through the ribosome. In addition to allowing interpretation of a posttranslocational state, eEF2 binding to the ribosome stabilizes the rpP0-P1-P2 (L7/L12 in bacteria) stalk. The dynamic motion of this region is likely important for ribosomal function, and we characterize the interactions of the stalk base with eEF2. Similarly, the L1 stalk is disordered in the crystal structures of 70S ribosomes, and its movement from an open to a closed state has been implicated in the function of protein synthesis using cryo-EM maps of various states (Valle et al., 2003). Our cryo-EM reconstruction contains a P/E hybrid state tRNA that interacts with the L1 stalk and stabilizes it in the closed position.

RESULTS AND DISCUSSION

Modeling the 80S Ribosome

In constructing an atomic model for the 80S ribosome from the cryo-EM map, we take advantage of X-ray structures solved for eubacterial ribosomes and their extensive sequence similarity with the *T. lanuginosus* ribosome. Because the resolution of the reconstructed density map is not sufficient to define the positions of side chains, the molecular structure we present should be seen as representative of an ensemble of atomic structures consistent with the experimental density, existing structural knowledge, and rules of stereochemistry.

To facilitate model building and docking, the density attributed to rRNA versus that ascribed to rps was computationally segmented by density thresholding, as described elsewhere (Spahn et al., 2000) (see Figure S1 available online). For each ribosomal subunit, the segmented RNA and protein maps were used to generate starting models of either all rRNA or all rps. Once completed, the rRNA and protein models were combined, and energy was minimized for the 40S and 60S subunits individually and then jointly within the entire 80S cryo-EM density. In all, 15 ribosomal proteins in the small subunit and 29 ribosomal proteins in the large subunit were modeled into the 80S ribosome on the basis of homology modeling (Table 1; Experimental Procedures). In addition, ribosomal proteins rpP0, RACK1, and rpL30e, which are unique to eukaryotes, were modeled and docked into the cryo-EM map according to the atomic structures of apo-proteins. Finally, we included the localization of rpS19e, another protein unique to 80S ribosomes, which has been linked to a congenital disease in humans called Diamond-Blackfan anemia (DBA).

40S Expansion Segments

ES6 is the largest ES, averaging 250 nt in eukaryotes (Neefs and De Wachter, 1990). Residing in the central domain of 18S, ES6

is composed of two halves: a 3′ half exhibiting conserved sequence identity among eukaryotes and a 5′ half with sequence variability that is absent in eubacterial and archaeal 16S rRNA (Cannone et al., 2002). Despite its presence in all eukaryotic 18S rRNA, the secondary structure interactions are difficult to predict, and ES6 is depicted as a large, disordered region in most secondary structure predictions. Recent progress based on phylogenetic comparative analysis, however, has been able to better predict the secondary structure of this large ES (Alkemar and Nygard, 2006; Wuyts et al., 2002). Comparative analysis of over 6000 species concluded that the predicted structure of ES6 is relatively conserved in most eukaryotes and that it likely consists of five hairpins and one internal helix (Alkemar and Nygard, 2006) (Figure 1A). Further support for the double- and single-stranded regions of this predicted structure was provided by chemical and enzymatic digestions of ES6 in wheat, yeast, and mouse ribosomes (Alkemar and Nygard, 2006). The density assigned to ES6 in our cryo-EM map further supports such secondary structure predictions, allowing us to assemble the first three-dimensional model of this complex segment of rRNA.

The entire region of ES6 is surrounded by density assigned to unidentified ribosomal proteins (Figure 2A), which likely assist in proper folding and assembly of the ES. The density of our map supports the notion that ES6 forms five hairpins in yeast 18S rRNA. The two 5′ helices, A and B, of ES6 (using the nomenclature of Alkemar and Nygard (2006)) display clear density that fits the two hairpins. Our data demonstrate that hairpin A of ES6 in yeast is conserved structurally with the entire sequence of ES6 in *E. coli* 16S rRNA. Previous structural analyses of a 15 Å cryo-EM map determined two masses of density that were attributed to ES6 (Spahn et al., 2001). We show that these two masses account for only two of the hairpins in ES6: hairpin E (see below) is accommodated by the mass of density that runs parallel to the length of the 40S subunit, whereas hairpin B is the mass running perpendicularly to the body of the 40S subunit and projecting into solution on the solvent side. Except at its base, hairpin B of ES6 does not interact with any other component of the ribosome. Hairpin C is a short hairpin that is largely disordered in our cryo-EM map. It appears, however, that this region of ES6 interacts intimately with a large region of unidentified ribosomal protein mass near the base of hairpin B (Figures 1B and 2A). Like hairpin C, hairpin D forms a short hairpin structure; however, density attributed to it is readily resolved as duplex RNA. This helix extends away from the 40S subunit between hairpins B and D in the 18S rRNA, interacting weakly with the mass of unidentified ribosomal protein near hairpin A.

Hairpin E of ES6 forms a long, bulbous structure on the 40S subunit, which is comparable in length to hairpin B. At the base of hairpin E, there is an interaction with a long C-terminal helix of rpL19e, which projects away from the 60S subunit and toward the 40S subunit (Figure 2B). Density, particularly near the stem loop of hairpin E, suggests that bases in this region of ES6 are interacting with other regions of the 40S subunit. Density near the end of this hairpin indicates an interaction with a long α helix, which belongs to a 40S ribosomal protein that is unidentified in our homology model. A second, intramolecular interaction is seen about two-thirds down the length of hairpin E, where density reveals a strong contact between hairpin E of ES6 and ES3 (Figure 2B). This interaction is particularly interesting because it has been implicated, but not yet shown, that nucleotides from ES3 form a tertiary interaction with those from ES6 in eukaryotes (Alkemar and Nygard, 2006). Our structure demonstrates that such a tertiary interaction likely exists.

Although the tertiary interaction with ES6 involves the stem of ES3, the apical stem loop of ES3 forms a strong contact with the tip of helix 9 of 18S rRNA. Density from ES3 has been localized to the left foot in a 15 Å cryo-EM reconstruction of the 80S ribosome from yeast (Spahn et al., 2001). Helix 9 in 18S rRNA is about 12 nt longer in yeast than in *E. coli*, supporting the extra density in this region of our cryo-EM reconstruction. On the basis of the improved resolution of the current reconstruction, we can now identify both apical stem loops of ES3 *and* helix 9 in the 18S rRNA that contribute to formation of the left foot of the 40S subunit (Figure 2).

ES7 is an extension of helix 26 in 16S rRNA, residing on the back of the shoulder of the 40S subunit. This extended helix is anchored to helix 23 of 18S rRNA in the body of the 40S subunit by a large mass of density, which belongs to unidentified ribosomal proteins, located between ES7 and rpS14 (rpS11p). Zero-length UV cross-linking of mRNA in 80S initiation complexes suggests that eukaryotic initiation factor 3 may bind near ES7 on the 40S subunit (Pisarev et al., 2008). Because the initiation of protein synthesis is particularly more complex in eukaryotes than in bacteria (more than 13 initiation factors involved versus three in bacteria), ES7 may have a role in recruiting or stabilizing some of the eukaryotic-specific initiation factors.

ES9 is located at the top of the head of the 40S subunit, where it interacts with protein rpS19e, which is discussed below. Finally, ES12 is an extension of helix 44 in 18S rRNA and forms the right foot of the 40S subunit, as previously pointed out (Spahn et al., 2001).

40S Ribosomal Proteins

In addition to comprising part of the 80S ribosome, eukaryotic ribosomal proteins have been shown to be involved in endonucleolytic cleavage events required for maturation of pre-40S and pre-60S particles in the nucleus and the cytoplasm of eukaryotic cells (Choesmel et al., 2007; Leger-Silvestre et al., 2004). Disruption of these events results in various human diseases, such as DBA. Mutations to ribosomal proteins leading to DBA were first identified in the gene coding rpS19e, but have since been associated with several other ribosomal proteins (Robledo et al., 2008). Although the mechanism of pathogenesis of DBA remains unclear, recent evidence points toward defects in ribosome biogenesis as being responsible (Gregory et al., 2007; Idol et al., 2007; Leger-Silvestre et al., 2004). Mutations in several other ribosomal proteins from both subunits, such as rpS17e, rpS24e, rpL5 (rpL18p), and rpL11 (rpL5p) (Robledo et al., 2008 and references therein), have been implicated to cause DBA, most likely through disruptions of ribosome biogenesis.

Ribosome biogenesis follows a conserved pathway that has been extensively studied in yeast. In this pathway, a preribosomal RNA transcript is processed to form the 18S ribosomal RNA (rRNA) of the 40S subunit, as well as the 25S/28S and 5.8S rRNA components of the 60S subunit. In rpS19e-depleted

Table 1. Ribosomal Proteins

Small subunit—E, A, B (Eukaryotes, Archaea, Bacteria)		Yeast Residues	Model Residues
Eukaryote 40S (Yeast)	Bacterial 30S (*E. coli*)		
rpS3	rpS3p	1–240	6-193
rpS9	rpS4p	1–197	6–21; 56–163
rpS2	rpS5p	1–254	75–236
	rpS6p		
rpS5	rpS7p	1–225	40–225
rpS22	rpS8p	1–130	6–129
rpS16	rpS9p	1–143	6–143
rpS20	rpS10p	1–121	21–116
rpS14	rpS11p	1–137	9–133
rpS23	rpS12p	1–145	28–145
rpS18	rpS13p	1–146	16–145
rpS29	rpS14p	1–56	5–54
rpS13	rpS15p	1–151	66–149
	rpS16p		
rpS11	rpS17p	1–156	67–146
	rpS18p		
rpS15	rpS19p	1–142	47–117
	rpS20p		
rpS0	rpS2p	1–252	9–201
	rpS21p		
Small subunit - E, A only			
rpS1e			
rpS4e			
rpS6e			
rpS8e			
rpS17e			
rpS19e		1–144	4–144
rpS24e			
rpS25e			
rpS26e			
rpS31e			
rpS27e			
rpS27e			
rpS28e			
rpS30e			
Small subunit—E only			
rpS7e			
rpS10e			
rpS12e			
rpS21e			
RACK1		1–319	7–319
Large subunit—E,A,B			
Eukaryote 60S (yeast)	Bacterial 60S (E. coli)		
rpL1	rpL1p	1–217	4–216

Table 1. *Continued*

Small subunit—E, A, B (Eukaryotes, Archaea, Bacteria)		Yeast Residues	Model Residues
	rpL25p		
rpL2	rpL2p	1–254	2–244
rpL3	rpL3p	1–387	5–366
rpL4	rpL4p	1–362	5–261
rpL11	rpL5p	1–174	2–152
rpL9	rpL6p	1–191	6–184
	rpL9p		
rpL16	rpL13p	1–199	7–184
rpL23	rpL14p	1–137	7–137
rpL28	rpL15p	1–149	6–147
	rpL16p		
	rpL17p		
rpL5	rpL18p	1–297	11–247
	rpL19p		
	rpL20p		
	rpL21p		
rpL17	rpL22p	1–184	4–153
rpL25	rpL23p	1–142	60–139
rpL26	rpL24p	1–127	9–124
	rpL25p		
	rpL27p		
	rpL28p		
rpL35	rpL29p	1–120	2–96
rpL7	rpL30p	1–244	10–55; 83–244
	rpL31p		
	rpL32p		
	rpL33p		
	rpL34p		
	rpL35p		
	rpL36p		
rpL12	rpL11p	1–165	7–144
rpP0	rpL10p	1–312	3–120
Large subunit—E, A only			
Eukaryote 60S (yeast)	Archaeal 50S (*H. marismortui*)		
rpL8e	rpL7Ae/rpL7a	1–256	107–219
rpL10e	rpL10e	1–221	4–168
rpL13e	rpL13e		
rpL14e	rpL14e		
rpL15e	rpL15e	1–204	2–193
rpL18e	rpL18e	1–186	21–141
rpL19e	rpL19e	1–189	2–177
rpL21e	rpL21e	1–160	2–117
rpL24e	rpL24e	1–155	8–52
rpL30e	rpL30e	1–105	8–105
rpL31e	rpL31e	1–113	9–87
rpL32e	rpL32e	1–130	14–122
rpL33e	rpL35Ae		

Table 1. *Continued*

Small subunit—E, A, B (Eukaryotes, Archaea, Bacteria)		Yeast Residues	Model Residues
Eukaryote 60S (Yeast)	Bacterial 50S (*E. coli*)		
rpL34e	rpL34e		
rpL37e	rpL37e	1–88	2–53
rpL39e	rpL39e	1–51	3–50
rpL40e	rpL40e		
rpL42e	rpL44e	1–106	2–93
rpL43e	rpL37Ae	1–92	11–82
rpL38e	rpL38e		
rpL41e	rpL41e	1–144	4–144
Large subunit—E only			
rpL6e			
rpL20e			
rpL22e			
rpL27e			
rpL29e			
rpL36e			

cells, defective cleavage leads to accumulation of novel 21S and 20S nuclear pre-rRNA molecules (Leger-Silvestre et al., 2004). This depletion of rpS19e, as well as mutations to rpS19e that cause DBA in humans, severely affects the production of 40S ribosomal subunits in yeast and human cells (Idol et al., 2007; Leger-Silvestre et al., 2004). Despite its central role in ribosome biogenesis and architecture, the location of rpS19e in the 80S ribosome has remained obscure.

The crystal structure of rpS19e from the archaeon, *Pyrococcus abyssi*, has recently been solved to atomic resolution (Gregory et al., 2007). We have located rpS19e in the density of our cryo-EM map using a motif search program (Rath et al., 2003) and constructed a homology model of yeast rpS19e based on the *P. abyssi* crystal structure. Protein rpS19e is located at the top of the "head" of the 40S subunit intercalated between helix 41 and ES9 of 18S rRNA (Figure 3). The location where rpS19e was found in the cryo-EM density map produces an excellent fit (cross-correlation coefficient = 0.79), which is corroborated by biochemical data. For example, rpS19e forms chemical cross-links with ribosomal proteins rpS5 (rpS7p), rpS16 (rpS9p), and rpS18 (rpS13p) in the 40S ribosomal subunit (Terao et al., 1980; Tolan and Traut, 1981; Uchiumi et al., 1981; Yeh et al., 1986). All three of these protein orthologs have been modeled into our atomic model of the 80S ribosome and are located in the head of the 40S subunit, proximal to the location of rpS19e presented here (Figure 1B; Figure S2). In addition, immunolabeling experiments have identified the general location of rpS19e to be in the head region of the 40S subunit (Lutsch et al., 1990). Finally, genetic evidence suggests that the nuclear protein Nep1p binds a specific sequence near helix 41 of 18S rRNA and supports the association of rpS19e to preribosomal particles in this vicinity (Buchhaupt et al., 2006). Our localization of rpS19e indicates that it does, in fact, interact closely with helix 41 of the 18S rRNA in the 40S subunit (Figure 3; Figure S2A). Interestingly, loops in our atomic model that are closest to helix 41 of the rRNA are disordered in the crystal structure of *P. abyssi* rpS19e, supporting a role in rRNA binding. The C-terminal helix of rpS19e contacts ES9 of the 18S rRNA in our model.

The location of rpS19e indicates that it may assist in the folding of rRNA in the head region of the pre-40S particle. This late processing step in the nucleus may provide a contact for nonribosomal factors required for subsequent maturation and nuclear export of the pre-40S particles. Indeed, rpS19e depletion has been shown to abrogate the binding of nonribosomal proteins, such as Tsr1, Rio2, and Enp1, that are required for subsequent cleavage events of pre-18S rRNA (Leger-Silvestre et al., 2004). Further support for this hypothesis comes from the fact that other ribosomal proteins, such as rpS5 (rpS7p) and rpS18 (rpS13p), which also reside in the head region of 40S subunits near rpS19e, display similar malfunctions in ribosome assembly when depleted in yeast cells (Ferreira-Cerca et al., 2005). Further studies will be required to fully understand the role of rpS19e in eukaryotic ribosome biogenesis; however, the location of rpS19e in the mature 40S particle identified here provides a basis for its interactions with rRNA and neighboring proteins in the eukaryotic ribosome.

Interestingly, it has recently been shown that mutations or disruptions of genes that encode ribosomal proteins are also implicated in other pediatric bone marrow failure syndromes in humans, including Shwachman-Diamond syndrome, Dyskeratosis congenita, cartilage-hair hypoplasia, and $5q^-$ syndrome (Ebert et al., 2008 and references within). Most notably, mutations to the genes that encode the ribosomal proteins rpS17e (Cmejla et al., 2007) and rpS24e (Choesmel et al., 2008; Gazda et al., 2006) have also been associated with DBA. These findings support the hypothesis that a loss of ribosome function is the cause of such diseases. The location of point mutations to rpS19e that result in DBA suggest such genetic alterations either inhibit proper folding of rpS19e, or they reside in areas that are adjacent to rRNA and their mutation likely abrogates a chaperone-like folding function of rpS19e on the ribosomal subunit (Supplemental Data).

In the small subunit, the general locations of rpS17e (Lutsch et al., 1979) and rpS24e (Lutsch et al., 1990) have been crudely mapped using immuno-EM of negatively stained ribosomes from rat liver. Antibodies against rpS17e were seen to bind in the head region of the 40S subunit at a location that remained accessible in the 80S ribosome. Antibodies against rpS24e bound to a single region on the interface surface of the 40S subunit. Of seven proteins investigated, rpS24e was the only 40S protein inaccessible to immunolabeling in 80S ribosome preparations, further supporting its existence in the intersubunit space (Lutsch et al., 1990). There are only two masses of density on the 60S interface of the 40S subunit in our cryo-EM map that could account for the location of rpS24e. These densities are seen on either side of helix 44, near ES12, of 18S rRNA. The structures of rpS24e from *Thermoplasma acidophilum* (Jeon et al., 2006) and *P. abyssi* (Choesmel et al., 2008) have been solved by NMR spectroscopy and X-ray crystallography, respectively. However, we were unable to confidently dock either structure into our cryo-EM map. We attribute this failure to a combination of the relatively small size of the apo protein (<100 residues) and the limited resolution of our cryo-EM map.

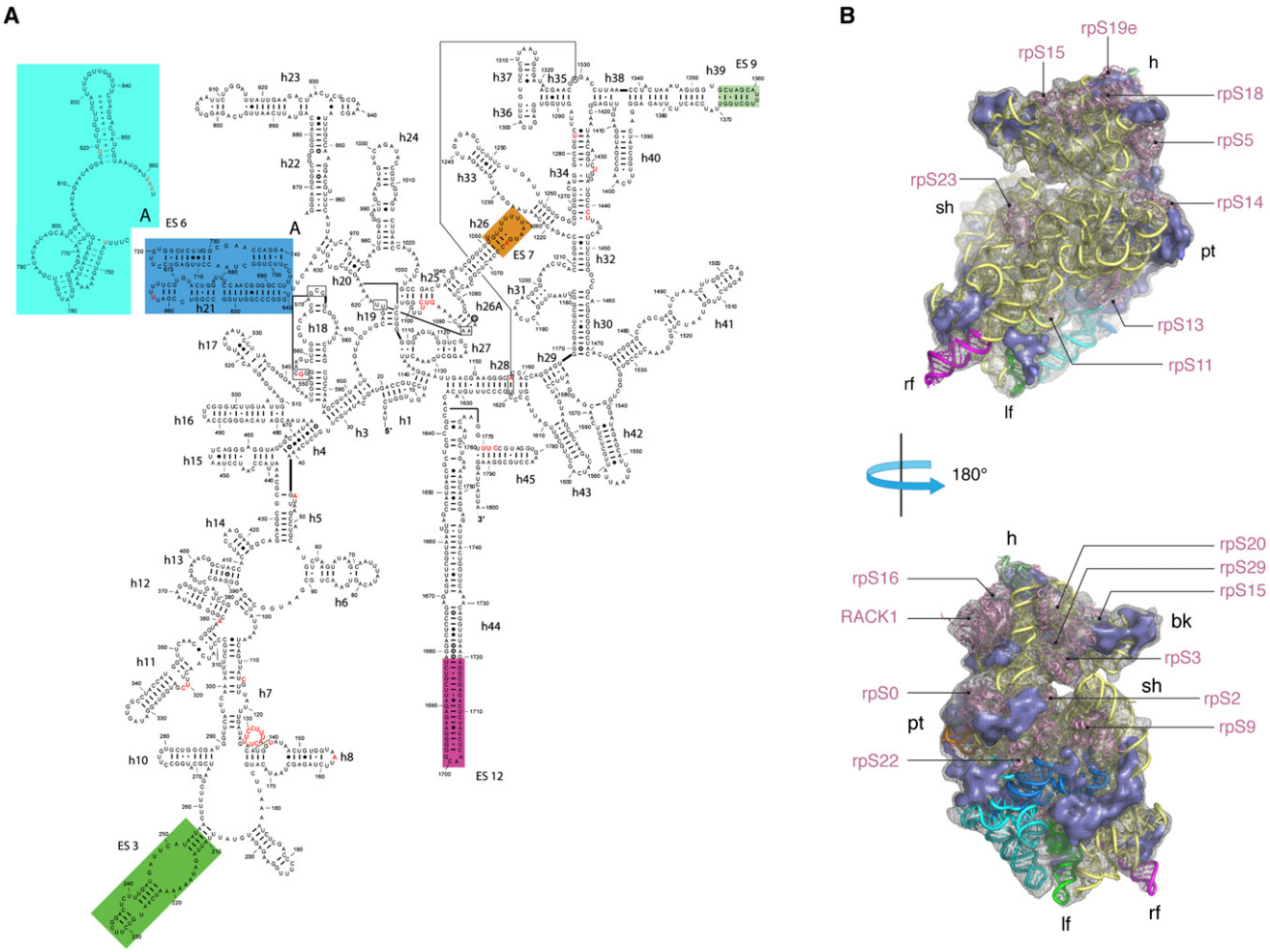

Figure 1. Structure of the Eukaryotic 40S Subunit

(A) Secondary structure prediction and rRNA helix numbering based primarily on that of the Comparative RNA Web Site of 18S rRNA of *S. cerevisiae* (http://www.rna.ccbb.utexas.edu). The secondary structure of ES6, however, is modeled based on that reported by Wuyts et al., (2002). Expansion segments are color-coded.

(B) The atomic model of the 40S subunit with ribosomal RNA represented as tubes, with expansion segments color-coded to match their secondary structure position presented in (A). Ribosomal proteins, with known homologs and placement, are shown as pink cartoons. Density accounting for proteins that are unique to eukaryotes with no structural homolog available or that have not been localized within the context of the ribosome is represented in slate. The boundary of the cryo-EM density of the 40S subunit is shown as a gray mesh.

RACK1

The receptor for activated C-kinase 1 (RACK1) is a WD40 repeat scaffold protein that is highly conserved in eukaryotic ribosomes and functions in a wide range of physiological processes, including ribosome assembly and activation, as well as gene transcription and translation (Nilsson et al., 2004). RACK1 has been localized on the back of the head of the 40S subunit by biochemical and cryo-EM techniques (Sengupta et al., 2004). Homology models of RACK1 have been constructed on the basis of the seven β-propeller architecture of the β subunit of heterotrimeric G proteins, a protein that shares ~25% sequence identity to RACK1 (Chandramouli et al., 2008). Recently, however, the X-ray structures of RACK1 (RACK1A) from *Arabidopsis thaliana* (Ullah et al., 2008) and from *S. cerevisiae* (Coyle et al., 2009) were solved, revealing differences between the β subunit of G proteins and RACK1.

Interactions between RACK1 and the 40S subunit are orchestrated primarily through blade 1 in RACK1 (Figure 4A). There is a weak interaction between the base of helix 40 in 18S rRNA and the first loop in blade one of RACK1 and a strong interaction between the base of helix 39 in 18S rRNA and the bottom of loops in blades 1 and 2 of RACK1. Genetic mutations revealed patches of basic residues within blade 1 that play a large role in RACK1's affinity for rRNA in the ribosome (Coyle et al., 2009). A rather strong interaction between rpS16 (rpS9p) of the 40S subunit and the middle of the β sheet in blade 1 of RACK1 is also apparent from the cryo-EM density. There is density attributed to an unidentified ribosomal protein in our model that is immediately adjacent to the bottom of the propeller in RACK1, which may be important for recruitment and binding of RACK1 to the 40S subunit. This interaction is orchestrated through the knoblike structure protruding between blades 6

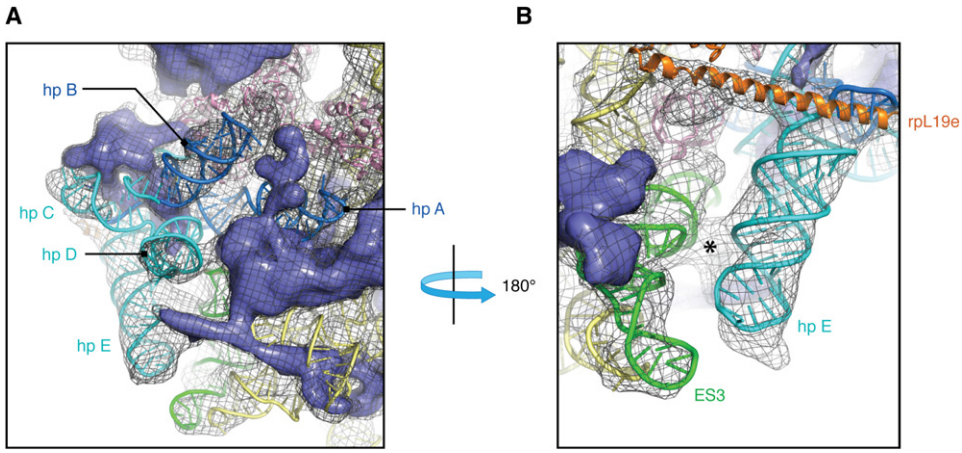

Figure 2. The Quaternary Structure of ES6

(A) Solvent view of the 40S subunit. ES6 is represented by two 5′ hairpins (A and B; blue tubes) followed by three hairpins on the 3′ end (C-E; cyan tubes). All five hairpins interact intimately with unknown proteins (slate density) on the solvent side of the 40S subunit.

(B) Interaction of ES3 (green) with hairpin E of ES6 (cyan tube) as viewed from the intersubunit space of the ribosome. Strong density in the EM map (marked with an asterisk) supports phylogenic data indicating that base pairing interaction exists between the two ESs. Prominent density seen for the C-terminal helix of rpL19e, reaching over from the 60S subunit, allowed the modeling of this region in our structure. This helix makes extensive interactions with ES6, particularly at the base of hairpin E.

and 7 of RACK1, which also comes into proximity of helix 40 in the 18S rRNA (Figure 4A). The knoblike protrusion does not appear to be involved in ribosomal binding, however, because its deletion has no adverse affects on ribosomal binding in vivo (Coyle et al., 2009).

RACK1 density is subdivided into two distinct halves in our cryo-EM map, which correspond to blades 1–3 and 4–7 (Figure 4B). Because RACK1 is known to interact with more than 80 different proteins (Sklan et al., 2006), it is conceivable that the two halves behave as docking stations for separate proteins to bind simultaneously, thus functioning as a chaperone responsible for bringing various proteins into proximity of one another. This notion is supported by the location of conserved residues throughout eukaryotic species that are within each half of the RACK1 propeller (Ullah et al., 2008).

RpS25e

Radiolabeled globin mRNA in the 80S initiation complex from rabbit reticulocytes has been shown to cross-link to rpS7e, rpS10e, rpS25e, rpS29 (rpS14p), and rpL5 (rpL18p) (Takahashi et al., 2005). Of these proteins, rpS5 (rpS7p) and rpL5 (rpL18p) have homologous proteins in eubacterial ribosomes and are indeed located along the mRNA channel. RpS25e does not have a bacterial homolog but cross-links directly to rpS5 (rpS7p) (Takahashi et al., 2005). On the basis of these data, there are two possible locations for rpS25e in our cryo-EM density. One potential density resides between protein rpS5 (rpS7p) and rpS18 (rpS13p), on top of the helix between helices 41 and 42 in 18S rRNA. A second possibility is near the exit channel, on the neck of the 40S subunit: interacting with the opposite side of rpS5 (rpS7p) and with helix 43 of 18S rRNA (Figure S3).

RpS26e and RpS28e

Site-directed, zero-length, UV cross-linking of mRNA identified four ribosomal proteins on the 40S subunit near the E-site of the 80S initiation complex. Two of these proteins, rpS5 (rpS7p) and rpS14 (rpS11p), have bacterial homologs and reside in the head and platform on the body of the 40S subunit, respectively. The two proteins that are eukaryote-specific and also cross-link to mRNA near the E-site of the ribosome are rpS26e and rpS28e. On the basis of our cryo-EM reconstruction, we hypothesize that

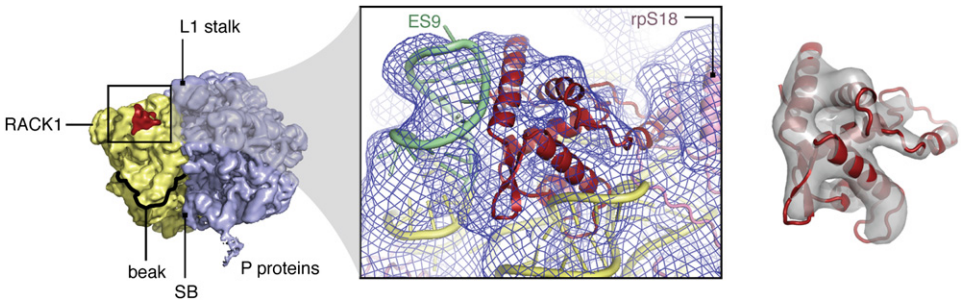

Figure 3. Localization of rpS19e in the 40S Ribosomal Subunit

rpS19e (red) localizes to the top of the head of the small subunit. Homology model of rpS19e crystal structure displayed within transparent cryo-EM density. Coloring is the same as in Figure 1.

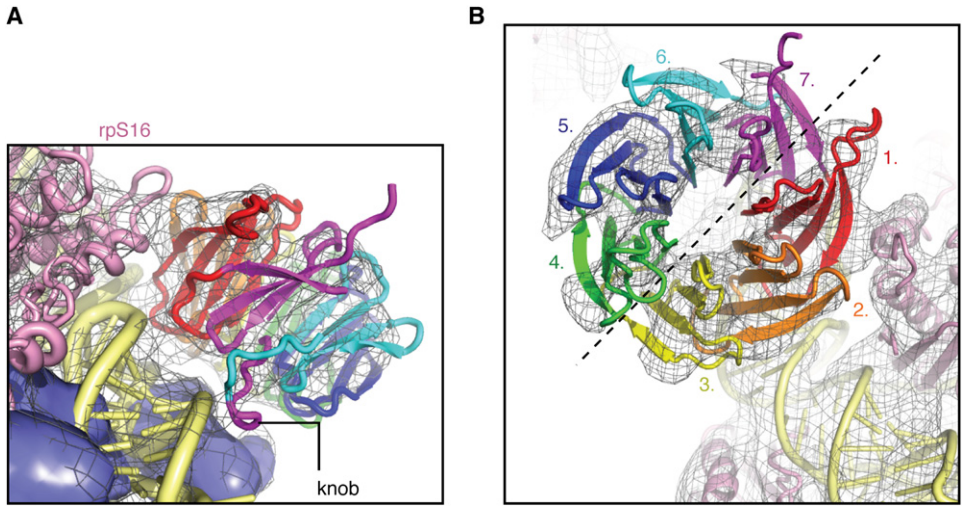

Figure 4. Interactions of RACK1 with the 40S Subunit

(A) Side view of RACK1 with its interaction to the 40S subunit. An unknown protein resides near RACK1, but the density between the two moieties is rather weak. Primary interactions between RACK1 and the 40S subunit are coordinated through 18S rRNA and rpS16 (rpS9p).

(B) Top view of RACK1 with the seven propellers color-coded. EM density indicates that primary interactions with the 40S subunit are coordinated via blade 1 of RACK1. Density of the RACK1 structure further indicates the molecule to have two distinct hemicycles with one half formed by tight interactions between blades 1–3 and a second half from blades 4–7.

both these proteins reside in the mass of unassigned density that resides on the platform of the 40S subunit surrounding rpS14 (rpS11p) and near the E site of the ribosome.

60S Expansion Segments

Expansion segments are found in all six domains of 5.8/25S rRNA of the large subunit (Figure 5); 5.8S rRNA contains one entire ES as well as part of a second ES. ES4 is a hairpin rRNA formed by the 3′ segment of 5.8S rRNA and the 5′ terminus of 25S rRNA (Figure 5; Figure S4A). This ES displays weak interactions with rpL8 (rpL7Ae) and an unassigned α helix. ES3 is formed entirely of 5.8S rRNA and resides between ES19 of 25S rRNA and rpL25 (rpL23p). The tip of ES3 makes a weak contact with helix 54 of 25S rRNA (Figure S4B).

The remaining expansion segments of the large subunit reside within 25S rRNA. A detailed description of the structure and architecture of these expansion segments is presented in Figure S4. However, some novel findings regarding expansion segments residing in the 60S subunit are presented here.

Because of the proximity of their points of origin and the limited resolution of previous cryo-EM reconstructions of fungal 80S ribosomes, ES7 and ES39 have been jointly assigned to the same general density on the back of the 60S subunit (Nilsson et al., 2007; Spahn et al., 2001). We are now able to distinguish these two expansion segments in our cryo-EM map as separate entities. Part of the difficulty in resolving these two rRNA regions at lower resolution must be attributed to their tight association with ribosomal proteins that have no homologs from the available X-ray atomic model. ES7 folds into two primary stem loops separated by a helix and has a third, very short stem loop at its 3′ end (Figure 5). The 5′ stem loop originates near the base of ES39 but protrudes away from the 60S subunit in the opposite direction. This duplex has slightly different conformations in the cryo-EM reconstructions of 80S ribosomes from *S. cerevisiae* and *T. lanuginosus*. In *S. cerevisiae*, this ES is seen jutting out into solution giving an "open" appearance, whereas it appears closer to the core of the ribosome in *T. lanuginosus*, which is best described as being "closed" (Nilsson et al., 2007). In our reconstruction, we can see that the "closed" conformation arises from the presence of an unidentified ribosomal protein anchoring this duplex of rRNA to the body of the 60S subunit (Figure 6A). Comparison to the density of an 11.7 Å reconstruction of the 80S ribosome from *S. cerevisiae* (Spahn et al., 2004) reveals the absence of this particular protein mass, which allows ES7 to adopt the "open" conformation (Figure 6B). It is possible that the protein is present in 80S ribosomes from *T. lanuginosus* and not those from *S. cerevisiae*, but the high sequence identity between the two species mentioned earlier makes that explanation unlikely. The genome of *T. lanuginosus*, including those genes that encode ribosomal proteins, has not been fully sequenced to resolve this issue. Nonetheless, two-dimensional gel electrophoresis of *T. lanuginosus* 80S ribosomes revealed the same number, and similar size, of ribosomal proteins in both subunits as those in *S. cerevisiae* (Wu et al., 1995). Therefore, we believe the protein density relates to a ribosomal protein conserved throughout fungi, which may become dissociated from the *S. cerevisiae* ribosome during purification while remaining bound to the *T. lanuginosus* ribosome. The tighter binding of this protein in *T. lanuginosus* might be required for the organism's increased thermostability.

The longest stem loop of ES7 resides about 70 Å away from ES39. Some unidentified ribosomal proteins surround this stem loop. Near the base of ES7 is a protein with an extraordinarily long α helix that stretches across the back of the 60S subunit. This α helix is approximately 80 Å in length and makes extensive contacts with the long stem loop, as well as the short 3′ stem loop structure, in ES7. On the basis of secondary structure predictions of unidentified, 60S ribosomal proteins, we predict

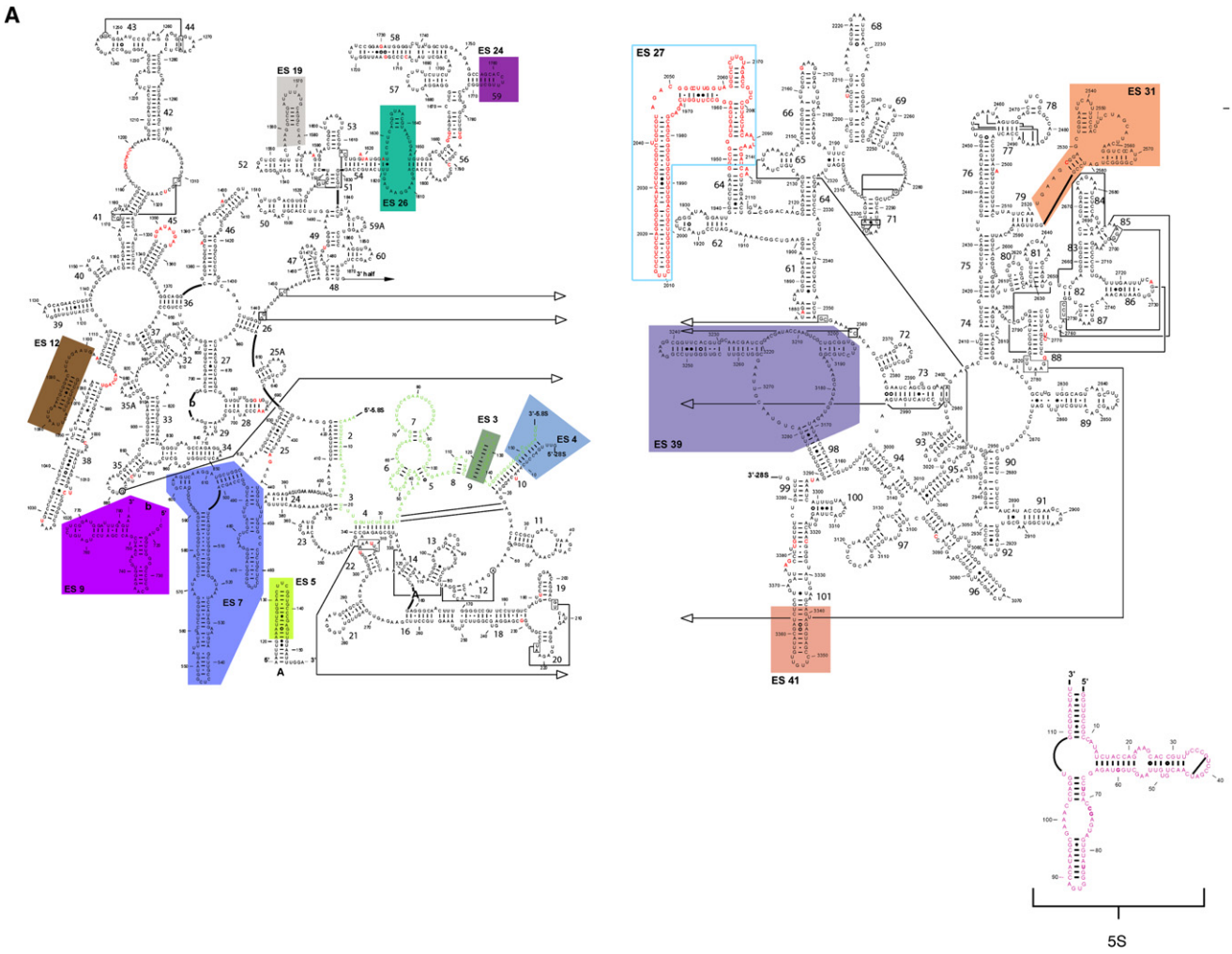

Figure 5. Structure of the Eukaryotic 60S Subunit

(A) Secondary structure prediction and rRNA helix numbering according to that of the Comparative RNA Web Site of rRNA of *S. cerevisiae* (http://www.rna.ccbb.utexas.edu). Expansion segments are color-coded.

(B) The atomic model of the 60S subunit with ribosomal RNA represented as tubes, with expansion segments color-coded to match their secondary structure position presented in (A). 5S and 5.8S rRNA are shown as magenta and green tubes, respectively. Ribosomal proteins, with known homologs and placement, are shown as orange cartoons. Density accounting for proteins that are unique to eukaryotes with no structural homolog available or that have not been localized within the context of the ribosome is represented in yellow. The boundary of the cryo-EM map of the 60S subunit is shown as a gray mesh.

that the long α helix is actually composed of two α helices from separate ribosomal proteins but that are in close proximity, giving the appearance of a single, very long α helix. The longer region of this density can be attributed to the N terminus of rpL7 (rpL30p), which is predicted to possess a C-terminal helix comprising about 55 residues in yeast and running along the longer hairpin of ES7 (Supplemental Data; Figure 5). The shorter portion of α-helical density likely belongs to the unidentified ribosomal protein, which anchors the 5′ hairpin of ES7 to 60S, as discussed previously.

60S Ribosomal Proteins: The Ribosomal Stalk Base, the P Proteins, and eEF2 Binding

The stalk base of the large ribosomal subunit comprises the universal GTPase-associated center (GAC) center of the ribosome. This region is composed of ribosomal RNA and proteins and is essential for recruitment, binding, and GTPase activity of ribosomal GTPases involved in initiation, decoding, translocation, and peptide release during protein synthesis (Frank et al., 2007). Ribosomal RNA from the large subunit, including the sarcin-ricin loop and protein rpP0 in eukaryotes (L10 in eubacteria) contribute to the architecture of the stalk base. A rpP0 homolog representing residues 3–120 (of 312 total) was modeled into our cryo-EM density, on the basis of the crystal structure of L10 from *Thermotoga maritima* (Diaconu et al., 2005) (Figure 7). The N-terminal portion of rpP0 is responsible for anchoring the stalk to a highly conserved region of 25S rRNA in the 60S subunit, while the CTD of rpP0 functions in anchoring the acidic P (P1 and P2) proteins to the stalk (Gudkov et al., 1980).

The eubacterial stalk base is composed of two or three L12 protein dimers that bind the CTD of L10 (also referred to as L7/L12) and form interactions with ribosomal factors, such as

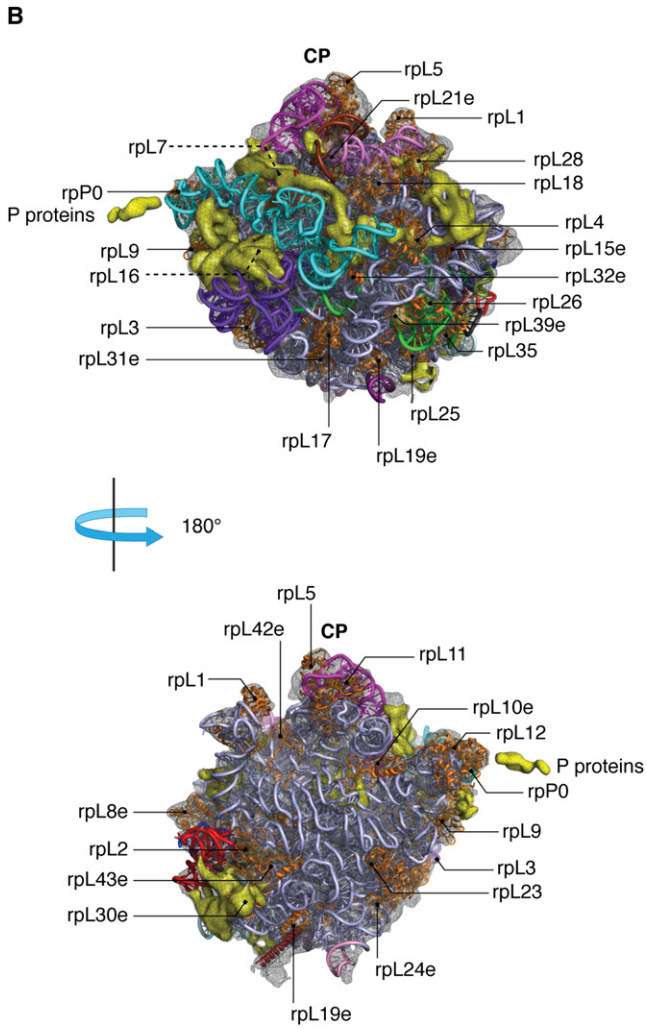

Figure 5. (*continued*).

EF-G, that are important for catalyzing protein synthesis (Datta et al., 2005; Diaconu et al., 2005). Two dimers of the acidic P1 and P2 proteins form a pentamer with the P0 protein, which constitutes the "L7/L12" stalk in eukaryotes. In yeast, slight alterations in sequence have yielded additional proteins, named P1A, P1B, P2A, and P2B, with dimers being formed by P1A-P2B and P1B-P2A polypeptides. The carboxyl domain of rpP0 (about 100 residues) is absent in the L10 orthologs and, in addition to sharing sequence homology with the P1 and P2 proteins, the CTD of P0 may also mimic the P1/P2 function. Reasoning for this hypothesis comes from the fact that elimination of all four P proteins (P1A, P2A, P1B, and P2B) has no lethal effect on cells (Remacha et al., 1995). In contrast, rpP0 is absolutely required, because disruption of its gene in yeast is lethal (Santos and Ballesta, 1994). It has been established in yeast that residues 199–230 and 231–258 of the CTD of rpP0 are responsible for binding the P1A-P2B and P1B-P2A dimers, respectively (Krokowski et al., 2006). The carboxyl end of rpP0 is structurally similar to the P1/P2 proteins and can function as the P1/P2 proteins in their absence (Santos and Ballesta, 1994). This highly conserved CTD interacts with eEF2 (Lalioti et al., 2002; Vard et al., 1997). As with the P1 and P2 acidic proteins, the CTD of rpP0 is predicted to comprise several helices. There is clear density that would account for the CTD of rpP0 contacting domain I of eEF2 in the structure presented here (Figure 7).

The N-terminal domains of three L7/L12 dimers from *T. maritima* have been solved in complex with L10 to atomic resolution (Diaconu et al., 2005). Diaconu et al. demonstrated that the X-ray crystal structure could be docked into an 18 Å cryo-EM map of the 70S ribosome from *E. coli* with high fidelity. In general, both the P and L12 proteins consist of an N- and C-terminal domain connected via an alanine-rich hinge. However, it has been argued that the L12 and P proteins are not structurally related, and there is no considerable similarity in sequence between the eukaryal/archaeal P proteins and the eubacterial L12 proteins (Grela et al., 2008). For these reasons, we have refrained from including the P1/P2 proteins in our model.

A series of α helices belonging to various ribosomal proteins of the 60S subunit, yet residing in regions with no corresponding structural homologs, were modeled into the cryo-EM density where the secondary structure was obvious and where the α helix was supported by secondary structure prediction using the Phyre server (Kelley and Sternberg, 2009). These helices were modeled for proteins rpL7 (rpL30p), rpL16 (rpL13p), rpL19e, rpL21e, and rpL35 (rpL29p) and, in addition to their architectural role in the 80S ribosome, possess other biological functions (Figure S5).

EXPERIMENTAL PROCEDURES

Protein-RNA Segmentation

The 8.9 Å cryo-EM map of the 80S ribosome from *T. lanuginosus* (accession number EMD-1345 in the 3D-EM database, EMBL-EBI; Taylor et al., 2007) was used for docking the 80S rRNA and rps. The electron-rich density attributed to rRNA was segmented from the density ascribed to ribosomal proteins taking account of the known volume ratio in the cryo-EM volume (Spahn et al., 2000).

RNA Modeling Methods

Overall strategy

We based most of the model for the yeast ribosomal RNA on the crystal structure of the *Escherichia coli* 70S ribosome (Berk et al., 2006; Schuwirth et al., 2005), making sequence changes, insertions and deletions as required. The ribosome of *Thermomyces lanuginosus*, which is the source of the cryo-EM map, shares over 85% sequence identity with that of *Saccharomyces cerevisiae*. Because the latter is much more widely studied than the former, we have used *S. cerevisiae* sequences for the model.

Alignment of RNA sequences

To determine homologous nucleotide positions between the *E. coli* and *S. cerevisiae* ribosomal RNA sequences, we used the ribosomal RNA alignments available on the Comparative RNA Website (CRW) (Cannone et al., 2002). This alignment results from a covariation analysis of a multiple sequence alignment using the ribosomal RNA sequences of hundreds of organisms. The CRW alignments are superior to those that might be inferred from traditional sequence alignment programs such as clustalw and BLAST.

Ribosomal RNA expansion segments

The secondary structure model of the *S. cerevisiae* rRNA from the CRW identifies expansion segments in the RNAs of both the large and small subunits. The secondary structures of some expansion segments were not available from the CRW. For these cases, we used MFOLD to predict the secondary structure (Zuker, 2003). For some large expansion segments (>200 nucleotides), such as ES6 in the small subunit, secondary structure has been proposed by using experimental methods (Alkemar and Nygard, 2006). Some

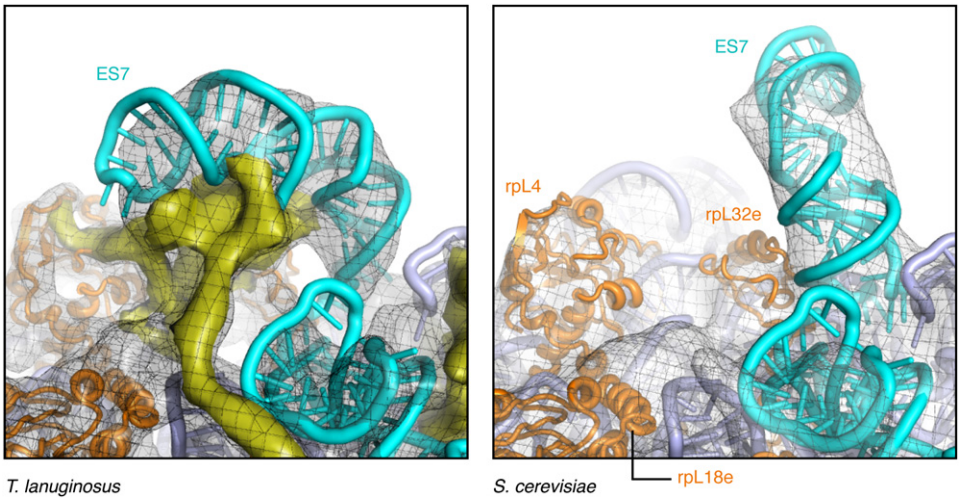

Figure 6. Protein Density Anchors ES7 to the Body of the 60S
ES7 of *T. lanuginosus* (left) is tethered to the body of the 60S subunit by an unidentified protein mass (yellow), presenting a "closed" conformation of ES7. This protein was absent in a previously reported structure of *S. cerevisiae* (right) (Spahn et al., 2001), leading to an "open" conformation of ES7.

small nucleotide segments within the ES of the large subunit (e.g. ES28 and ES39) lacked both experimental data and useable MFOLD predictions. We modeled these as single strands or loops.

Changing the* E. coli *homologous nucleotides to yeast

Given the sequence alignment (which specifies the correspondence between nucleotides in the yeast and *E. coli* rRNAs), single nucleotide changes were easily made in double-helical regions and in those single-stranded and loop regions where the yeast and *E. coli* sequences are of equal length. Nucleotide substitutions in these regions were made using the Biopolymer module of the Insight-II software package (Molecular Simulations, Inc., San Diego, CA). In those cases where insertions and/or deletions were required, we used loops with the desired number of bases from crystal structures, manually replacing the existing loop with the model loop, with manual adjustments to get reasonable fits prior to minimization. In most cases, some nucleotides needed to be substituted to fit the yeast sequence. Variable regions such as helix 33 (the beak region of SSU) contained many changes that needed extensive manual intervention.

Secondary to tertiary structure

For those ES with known secondary structures, the MCSYM program was used to generate tertiary structures (Major et al., 1991). When generating tertiary structures, MCSYM takes into account structural constraints such as pairing partners and the type of interaction between them, connection

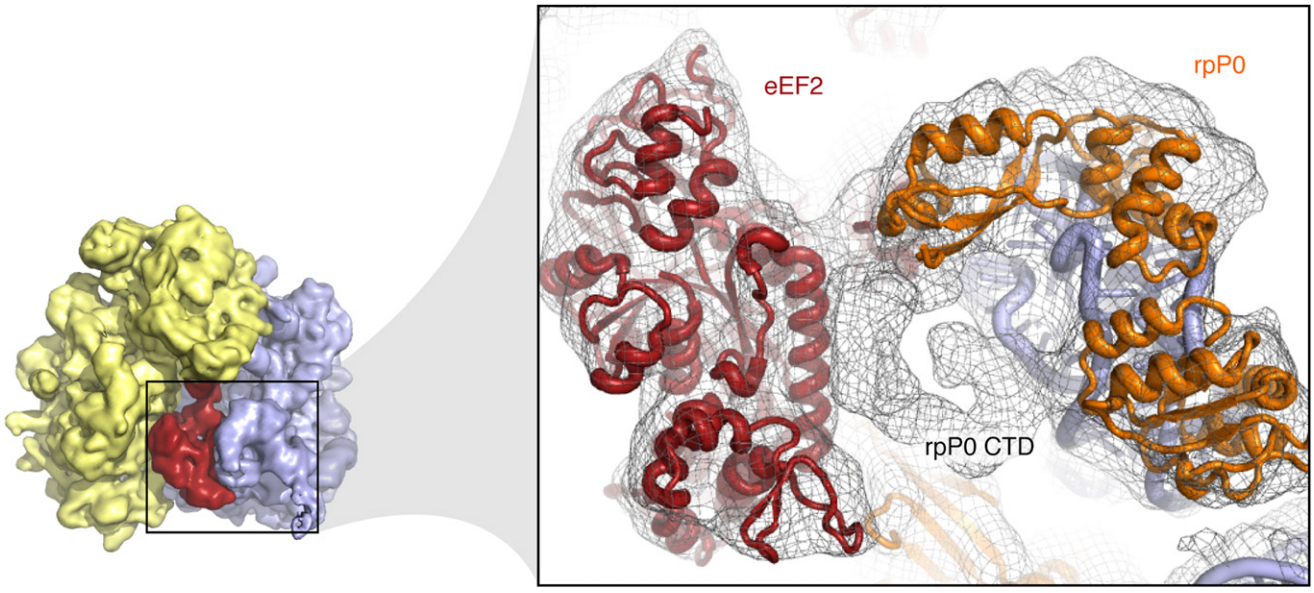

Figure 7. Interaction of rpP0 CTD with eEF2
The NTD of rpP0 was modeled into the density of our cryo-EM map, on the basis of the orthologous X-ray structure of L10 from *T. maritime* (Diaconu et al., 2005). Density for the CTD of rpP0 is clearly visible and indicative of α-helical secondary structure. The density further indicates a clear interaction between the rpP0 CTD and a long α-helix residing domain I of eEF2.

information between consecutive residues, and so on, as specified by the user. The program then searches through a structural database and gathers all the pairs that satisfy the constraints. Using the residues gathered, it generates a three-dimensional model. The structures generated by MCSYM are not always energetically favorable, because the models generated satisfy only the constraints provided and do not use energetic constraints, and MCSYM does not refine these initial models. Thus, the models obtained from MCSYM need to be analyzed and minimized using a molecular mechanics program. For this purpose, we used the Insight-II Discover module.

Fitting the RNA expansion segments into the density map

Chimera (Pettersen et al., 2004) was used to view the density map and perform manual fitting, roughly placing the models inside the appropriate regions of density. The fit-in-map module of Chimera was used to fit the rigid ribosomal RNA model into the density map. For those expansion segments consisting of multiple helical regions, the helices were separated at the loop regions, each helix was independently fitted into the density, and the helices were then reconnected. Helical regions are well ordered in the density maps, allowing us to fit those regions of our model with confidence.

Some expansion segments share the same cryo-EM density, especially in the large subunit, so careful modeling of the RNA was essential to avoid overlap. A combination of manual and automated methods was used in these efforts, and sometimes multiple iterations were required to generate satisfactory models.

Optimizing fit and RNA structure refinement

The initial fit to the cryo-EM density of *T. lanuginosus* was refined with a flexible fitting algorithm using the Emmental sub-package in the YUP.scx module of YUP (http://rumour.biology.gatech.edu/YammpWeb) (Tan et al., 2006). A Gaussian Network Model (GNM) represents the all-atom structure. The energy function contains terms for scoring the quality of the fit of the model to the density map, plus restraint energies for the GNM and volume exclusion terms. The optimization protocol uses simulated annealing with molecular dynamics. This optimization produced some steric clashes. These included a handful of ring penetrations, which we corrected manually, before minimization with NAMD (Phillips et al., 2005), using the AMBER force field (Cornell et al., 1995).

Comparative Protein Structure Modeling

There are 78 known proteins in the yeast ribosome, of which 44 have homologs available in archaeal and bacterial ribosome structures. We calculated models for these protein sequences using MODPIPE, an automated software pipeline for large-scale protein structure modeling (Eswar et al., 2003). MODPIPE relies on MODELER (Sali and Blundell, 1993) for its functionality and can calculate comparative models for a large number of sequences using different template structures and sequence-structure alignments. Sequence-structure matches were established using a variety of fold-assignment methods, including sequence-sequence (Smith and Waterman, 1981), profile-sequence (Altschul et al., 1997; Eswar et al., 2005a) and profile-profile alignments (Eswar et al., 2005b; Marti-Renom et al., 2004). Increased sensitivity of the search for known template structures was achieved by using an E-value threshold of 1.0. Ten models were calculated for each of the sequence-structure matches resulting in a reasonable degree of conformational sampling. The best scoring model for each alignment was then chosen using the DOPE score, a distance-dependent atomic statistical potential (Shen and Sali, 2006). Finally, all models generated for a given input sequence were evaluated for the correctness of the fold using a composite model quality criterion that included the coverage of the model, sequence identity of the sequence-structure alignment, the fraction of gaps in the alignment, the compactness of the model, and statistical potential Z-scores (Eramian et al., 2008; Melo et al., 2002; Shen and Sali, 2006). Only models that were assessed to have the correct fold were considered for docking into the EM density map.

Fitting of ribosomal proteins

Each of the models calculated with the procedure described above was rigidly fitted into the *T. lanuginosus* density map using Mod-EM (Topf et al., 2005). The initial position of each protein model was assigned by superposing its coordinates onto the corresponding coordinates of comparative models previously fitted into the 11.7 Å cryo-EM map of the yeast ribosome (pdb id: 1S1H: 40S subunit, 1S1I: 60S subunit; Spahn et al., 2004). The search for the best fit was then achieved by a local exhaustive exploration of Euler angles, guided by a cross-correlation coefficient (CCC) between the model and the density map (Topf et al., 2005). To further validate the best fits they were compared to the best local fits achieved using the Chimera fit-in-map module. For each protein, the best model was selected on the basis of a combination of the highest CCC and the lowest DOPE score (Eramian et al., 2008; Shen and Sali, 2006; Topf et al., 2005, 2006).

The next step was a visual inspection of all fitted models using Chimera (Goddard et al., 2005) in the context of the fitted rRNA chains. The inspection revealed some atom clashes (between two protein models or between a protein model and an rRNA chain) and some rigid parts (such as individual domains or groups of α helices) that were shifted or rotated relative to the density. For these cases, we performed flexible fitting to refine the models, using Flex-EM (Topf et al., 2008). Parts of the models without clear density were omitted from the initial models.

The ribosomal proteins were then added to the rRNA model using Insight-II and VMD. Because the proteins were modeled separately, we inspected the interface between each protein and RNA for any steric clashes. In most cases, only minor manual adjustments were needed, and each protein was then energy minimized along with the neighboring rRNA using Insight-II and NAMD for 1000 steps. To achieve the correct stereochemistry and eliminate stearic clashes, each subunit of the final combined protein and nucleic acid model was minimized for 200,000 steps using the steepest descent algorithm with NAMD, using the CHARMM force field. The final steric acceptability of the model was verified through extensive visual inspection of the final model, followed by a quantitative evaluation using the RCSB Protein Databank's Automatic Deposition and Inspection Tool (ADIT). The final CCC value of the entire model was 0.72.

ACCESSION NUMBERS

The atomic models of the 60S rRNA and rps have been deposited separately into the Protein Databank with accession codes 3JYX and 3JYW, respectively. The atomic model of the entire 40S subunit, including rRNA, rps, and P/E-tRNA, has been deposited into the Protein Databank with accession code 3JYV.

SUPPLEMENTAL DATA

Supplemental Data include one table, six figures, and Supplemental Experimental Procedures and can be found with this article online at http://www.cell.com/structure/supplemental/S0969-2126(09)00416-X.

ACKNOWLEDGMENTS

This work was supported by the Howard Hughes Medical Institute (support to J.F.), the Georgia Research Alliance (support to S.C.H.), the National Institutes of Health (grants R37 GM29169 and GM55440 to J.F.; grants R01 BM54762, U54 GM074945, PN2 EY016525, R01 BM083960, and NSF IIS-0705196 to A.S.; grant GM53827 to S.C.H.; and grant RR12255 to C.L. Brooks, III, Principal Investigator), and by an MRC Career Development Award (G0600084 to M.T.). We thank Bimal Rath for assistance with the motif search and Lila Rubenstein for assistance with the illustrations.

Received: July 21, 2009
Revised: September 16, 2009
Accepted: September 18, 2009
Published: December 8, 2009

REFERENCES

Alkemar, G., and Nygard, O. (2006). Probing the secondary structure of expansion segment ES6 in 18S ribosomal RNA. Biochemistry *45*, 8067–8078.

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. *25*, 3389–3402.

Berk, V., Zhang, W., Pai, R.D., and Cate, J.H. (2006). Structural basis for mRNA and tRNA positioning on the ribosome. Proc. Natl. Acad. Sci. USA *103*, 15830–15834.

 ©2009 Elsevier Ltd All rights reserved

Blaha, G., Stanley, R.E., and Steitz, T.A. (2009). Formation of the first peptide bond: the structure of EF-P bound to the 70S ribosome. Science *325*, 966–970.

Buchhaupt, M., Meyer, B., Kotter, P., and Entian, K.D. (2006). Genetic evidence for 18S rRNA binding and an Rps19p assembly function of yeast nucleolar protein Nep1p. Mol. Genet. Genomics *276*, 273–284.

Cannone, J.J., Subramanian, S., Schnare, M.N., Collett, J.R., D'Souza, L.M., Du, Y., Feng, B., Lin, N., Madabusi, L.V., Muller, K.M., et al. (2002). The comparative RNA web (CRW) site: an online database of comparative sequence and structure information for ribosomal, intron, and other RNAs. BMC Bioinformatics *3*, 2.

Chandramouli, P., Topf, M., Menetret, J.F., Eswar, N., Cannone, J.J., Gutell, R.R., Sali, A., and Akey, C.W. (2008). Structure of the mammalian 80S ribosome at 8.7 A resolution. Structure *16*, 535–548.

Choesmel, V., Bacqueville, D., Rouquette, J., Noaillac-Depeyre, J., Fribourg, S., Cretien, A., Leblanc, T., Tchernia, G., Da Costa, L., and Gleizes, P.E. (2007). Impaired ribosome biogenesis in Diamond-Blackfan anemia. Blood *109*, 1275–1283.

Choesmel, V., Fribourg, S., Aguissa-Toure, A.H., Pinaud, N., Legrand, P., Gazda, H.T., and Gleizes, P.E. (2008). Mutation of ribosomal protein RPS24 in Diamond-Blackfan anemia results in a ribosome biogenesis disorder. Hum. Mol. Genet. *17*, 1253–1263.

Cmejla, R., Cmejlova, J., Handrkova, H., Petrak, J., and Pospisilova, D. (2007). Ribosomal protein S17 gene (RPS17) is mutated in Diamond-Blackfan anemia. Hum. Mutat. *28*, 1178–1182.

Cornell, W.D., Cieplak, P., Bayly, C.I., Gould, I.R., Merz, K.M., Jr., Ferguson, D.M., Spellmeyer, D.C., Fox, T., Caldwell, J.W., and Kollman, P.A. (1995). A second generation force field for the simulation of proteins, nucleic acids and organic molecules. J. Am. Chem. Soc. *117*, 5179–5197.

Coyle, S.M., Gilbert, W.V., and Doudna, J.A. (2009). Direct link between RACK1 function and localization at the ribosome in vivo. Mol. Cell. Biol. *29*, 1626–1634.

Datta, P.P., Sharma, M.R., Qi, L., Frank, J., and Agrawal, R.K. (2005). Interaction of the G' domain of elongation factor G and the C-terminal domain of ribosomal protein L7/L12 during translocation as revealed by cryo-EM. Mol. Cell *20*, 723–731.

Diaconu, M., Kothe, U., Schlunzen, F., Fischer, N., Harms, J.M., Tonevitsky, A.G., Stark, H., Rodnina, M.V., and Wahl, M.C. (2005). Structural basis for the function of the ribosomal L7/12 stalk in factor binding and GTPase activation. Cell *121*, 991–1004.

Ebert, B.L., Pretz, J., Bosco, J., Chang, C.Y., Tamayo, P., Galili, N., Raza, A., Root, D.E., Attar, E., Ellis, S.R., and Golub, T.R. (2008). Identification of RPS14 as a 5q- syndrome gene by RNA interference screen. Nature *451*, 335–339.

Eramian, D., Eswar, N., Shen, M.Y., and Sali, A. (2008). How well can the accuracy of comparative protein structure models be predicted? Protein Sci. *17*, 1881–1893.

Eswar, N., John, B., Mirkovic, N., Fiser, A., Ilyin, V.A., Pieper, U., Stuart, A.C., Marti-Renom, M.A., Madhusudhan, M.S., Yerkovich, B., and Sali, A. (2003). Tools for comparative protein structure modeling and analysis. Nucleic Acids Res. *31*, 3375–3380.

Eswar, N., Madhusudhan, M.S., Marti-Renom, M., and Sali, A. (2005a). BUILD_PROFILE: A module for calculating sequence profiles in MODELLER (http://www.salilab.org/modeller).

Eswar, N., Madhusudhan, M.S., Marti-Renom, M., and Sali, A. (2005b). PROFILE_SCAN: A module for fold-assignment using profile-profile scanning in MODELLER (http://www.salilab.org/modeller).

Ferreira-Cerca, S., Poll, G., Gleizes, P.E., Tschochner, H., and Milkereit, P. (2005). Roles of eukaryotic ribosomal proteins in maturation and transport of pre-18S rRNA and ribosome function. Mol. Cell *20*, 263–275.

Frank, J., Gao, H., Sengupta, J., Gao, N., and Taylor, D.J. (2007). The process of mRNA-tRNA translocation. Proc. Natl. Acad. Sci. USA *104*, 19671–19678.

Gazda, H.T., Grabowska, A., Merida-Long, L.B., Latawiec, E., Schneider, H.E., Lipton, J.M., Vlachos, A., Atsidaftos, E., Ball, S.E., Orfali, K.A., et al. (2006). Ribosomal protein S24 gene is mutated in Diamond-Blackfan anemia. Am. J. Hum. Genet. *79*, 1110–1118.

Gerbi, S.A. (1996). Expansion segments: regions of variable size that interrupt the universal core secondary structure of ribosomal RNA. In Ribosomal RNA. Structure, Evolution, Processing, and Function in Protein Biosynthesis, R.A. Zimmerman and A.E. Dahlberg, eds. (Boca Raton, FL: CRC Press), pp. 71–87.

Goddard, T.D., Huang, C.C., and Ferrin, T.E. (2005). Software extensions to UCSF chimera for interactive visualization of large molecular assemblies. Structure *13*, 473–482.

Gregory, L.A., Aguissa-Toure, A.H., Pinaud, N., Legrand, P., Gleizes, P.E., and Fribourg, S. (2007). Molecular basis of Diamond-Blackfan anemia: structure and function analysis of RPS19. Nucleic Acids Res. *35*, 5913–5921.

Grela, P., Bernado, P., Svergun, D., Kwiatowski, J., Abramczyk, D., Grankowski, N., and Tchorzewski, M. (2008). Structural relationships among the ribosomal stalk proteins from the three domains of life. J. Mol. Evol. *67*, 154–167.

Gudkov, A.T., Tumanova, L.G., Gongadze, G.M., and Bushuev, V.N. (1980). Role of different regions of ribosomal proteins L7 and L10 in their complex formation and in the interaction with the ribosomal 50 S subunit. FEBS Lett. *109*, 34–38.

Gutell, R.R., Weiser, B., Woese, C.R., and Noller, H.F. (1985). Comparative anatomy of 16-S-like ribosomal RNA. Prog. Nucleic Acid Res. Mol. Biol. *32*, 155–216.

Idol, R.A., Robledo, S., Du, H.Y., Crimmins, D.L., Wilson, D.B., Ladenson, J.H., Bessler, M., and Mason, P.J. (2007). Cells depleted for RPS19, a protein associated with Diamond Blackfan Anemia, show defects in 18S ribosomal RNA synthesis and small ribosomal subunit production. Blood Cells Mol. Dis. *39*, 35–43.

Jeon, B.Y., Jung, J., Kim, D.W., Yee, A., Arrowsmith, C.H., and Lee, W. (2006). Solution structure of TA1092, a ribosomal protein S24e from *Thermoplasma acidophilum*. Proteins *64*, 1095–1097.

Kelley, L.A., and Sternberg, M.J. (2009). Protein structure prediction on the Web: a case study using the Phyre server. Nat. Protoc. *4*, 363–371.

Krokowski, D., Boguszewska, A., Abramczyk, D., Liljas, A., Tchorzewski, M., and Grankowski, N. (2006). Yeast ribosomal P0 protein has two separate binding sites for P1/P2 proteins. Mol. Microbiol. *60*, 386–400.

Lalioti, V.S., Perez-Fernandez, J., Remacha, M., and Ballesta, J.P. (2002). Characterization of interaction sites in the *Saccharomyces cerevisiae* ribosomal stalk components. Mol. Microbiol. *46*, 719–729.

Leger-Silvestre, I., Milkereit, P., Ferreira-Cerca, S., Saveanu, C., Rousselle, J.C., Choesmel, V., Guinefoleau, C., Gas, N., and Gleizes, P.E. (2004). The ribosomal protein Rps15p is required for nuclear exit of the 40S subunit precursors in yeast. EMBO J. *23*, 2336–2347.

Lutsch, G., Noll, F., Theise, H., Enzmann, G., and Bielka, H. (1979). Localization of proteins S1, S2, S16 and S23 on the surface of small subunits of rat liver ribosomes by immune electron microscopy. Mol. Gen. Genet. *176*, 281–291.

Lutsch, G., Stahl, J., Kargel, H.J., Noll, F., and Bielka, H. (1990). Immunoelectron microscopic studies on the location of ribosomal proteins on the surface of the 40S ribosomal subunit from rat liver. Eur. J. Cell Biol. *51*, 140–150.

Major, F., Turcotte, M., Gautheret, D., Lapalme, G., Fillion, E., and Cedergren, R. (1991). The combination of symbolic and numerical computation for three-dimensional modeling of RNA. Science *253*, 1255–1260.

Marti-Renom, M.A., Madhusudhan, M.S., and Sali, A. (2004). Alignment of protein sequences by their profiles. Protein Sci. *13*, 1071–1087.

Melo, F., Sanchez, R., and Sali, A. (2002). Statistical potentials for fold assessment. Protein Sci. *11*, 430–448.

Musters, W., Boon, K., van der Sande, C.A., van Heerikhuizen, H., and Planta, R.J. (1990). Functional analysis of transcribed spacers of yeast ribosomal DNA. EMBO J. *9*, 3989–3996.

Neefs, J.M., and De Wachter, R. (1990). A proposal for the secondary structure of a variable area of eukaryotic small ribosomal subunit RNA involving the existence of a pseudoknot. Nucleic Acids Res. *18*, 5695–5704.

Nilsson, J., Sengupta, J., Frank, J., and Nissen, P. (2004). Regulation of eukaryotic translation by the RACK1 protein: a platform for signalling molecules on the ribosome. EMBO Rep. *5*, 1137–1141.

Nilsson, J., Sengupta, J., Gursky, R., Nissen, P., and Frank, J. (2007). Comparison of fungal 80 S ribosomes by cryo-EM reveals diversity in structure and conformation of rRNA expansion segments. J. Mol. Biol. *369*, 429–438.

Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. (2004). UCSF Chimera–a visualization system for exploratory research and analysis. J. Comput. Chem. *25*, 1605–1612.

Phillips, J.C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R.D., Kale, L., and Schulten, K. (2005). Scalable molecular dynamics with NAMD. J. Comput. Chem. *26*, 1781–1802.

Pisarev, A.V., Kolupaeva, V.G., Yusupov, M.M., Hellen, C.U., and Pestova, T.V. (2008). Ribosomal position and contacts of mRNA in eukaryotic translation initiation complexes. EMBO J. *27*, 1609–1621.

Rath, B.K., Hegerl, R., Leith, A., Shaikh, T.R., Wagenknecht, T., and Frank, J. (2003). Fast 3D motif search of EM density maps using a locally normalized cross-correlation function. J. Struct. Biol. *144*, 95–103.

Remacha, M., Jimenez-Diaz, A., Santos, C., Briones, E., Zambrano, R., Rodriguez Gabriel, M.A., Guarinos, E., and Ballesta, J.P. (1995). Proteins P1, P2, and P0, components of the eukaryotic ribosome stalk. New structural and functional aspects. Biochem. Cell Biol. *73*, 959–968.

Robledo, S., Idol, R.A., Crimmins, D.L., Ladenson, J.H., Mason, P.J., and Bessler, M. (2008). The role of human ribosomal proteins in the maturation of rRNA and ribosome production. RNA *14*, 1918–1929.

Sali, A., and Blundell, T.L. (1993). Comparative protein modelling by satisfaction of spatial restraints. J. Mol. Biol. *234*, 779–815.

Santos, C., and Ballesta, J.P. (1994). Ribosomal protein P0, contrary to phosphoproteins P1 and P2, is required for ribosome activity and *Saccharomyces cerevisiae* viability. J. Biol. Chem. *269*, 15689–15696.

Schuwirth, B.S., Borovinskaya, M.A., Hau, C.W., Zhang, W., Vila-Sanjurjo, A., Holton, J.M., and Cate, J.H. (2005). Structures of the bacterial ribosome at 3.5 A resolution. Science *310*, 827–834.

Sengupta, J., Nilsson, J., Gursky, R., Spahn, C.M., Nissen, P., and Frank, J. (2004). Identification of the versatile scaffold protein RACK1 on the eukaryotic ribosome by cryo-EM. Nat. Struct. Mol. Biol. *11*, 957–962.

Shen, M.Y., and Sali, A. (2006). Statistical potential for assessment and prediction of protein structures. Protein Sci. *15*, 2507–2524.

Sklan, E.H., Podoly, E., and Soreq, H. (2006). RACK1 has the nerve to act: structure meets function in the nervous system. Prog. Neurobiol. *78*, 117–134.

Smith, T.F., and Waterman, M.S. (1981). Identification of common molecular subsequences. J. Mol. Biol. *147*, 195–197.

Spahn, C.M., Penczek, P.A., Leith, A., and Frank, J. (2000). A method for differentiating proteins from nucleic acids in intermediate-resolution density maps: cryo-electron microscopy defines the quaternary structure of the *Escherichia coli* 70S ribosome. Structure *8*, 937–948.

Spahn, C.M., Beckmann, R., Eswar, N., Penczek, P.A., Sali, A., Blobel, G., and Frank, J. (2001). Structure of the 80S ribosome from *Saccharomyces cerevisiae*–tRNA-ribosome and subunit-subunit interactions. Cell *107*, 373–386.

Spahn, C.M., Gomez-Lorenzo, M.G., Grassucci, R.A., Jorgensen, R., Andersen, G.R., Beckmann, R., Penczek, P.A., Ballesta, J.P., and Frank, J. (2004). Domain movements of elongation factor eEF2 and the eukaryotic 80S ribosome facilitate tRNA translocation. EMBO J. *23*, 1008–1019.

Sweeney, R., Chen, L., and Yao, M.C. (1994). An rRNA variable region has an evolutionarily conserved essential role despite sequence divergence. Mol. Cell. Biol. *14*, 4203–4215.

Takahashi, Y., Hirayama, S., and Odani, S. (2005). Ribosomal proteins cross-linked to the initiator AUG codon of a mRNA in the translation initiation complex by UV-irradiation. J Biochem *138*, 41–46.

Tan, R.K.Z., Petrov, A.S., and Harvey, S.C. (2006). YUP: A molecular simulation program for coarse-grained and multiscaled models. J. Chem. Theory Comput. *2*, 529–540.

Taylor, D.J., Nilsson, J., Merrill, A.R., Andersen, G.R., Nissen, P., and Frank, J. (2007). Structures of modified eEF2 80S ribosome complexes reveal the role of GTP hydrolysis in translocation. EMBO J. *26*, 2421–2431.

Terao, K., Uchiumi, T., Kobayashi, Y., and Ogata, K. (1980). Identification of neighbouring protein pairs in the rat liver 40-S ribosomal subunits cross-linked with dimethyl suberimidate. Biochim. Biophys. Acta *621*, 72–82.

Tolan, D.R., and Traut, R.R. (1981). Protein topography of the 40 S ribosomal subunit from rabbit reticulocytes shown by cross-linking with 2-iminothiolane. J. Biol. Chem. *256*, 10129–10136.

Topf, M., Baker, M.L., John, B., Chiu, W., and Sali, A. (2005). Structural characterization of components of protein assemblies by comparative modeling and electron cryo-microscopy. J. Struct. Biol. *149*, 191–203.

Topf, M., Baker, M.L., Marti-Renom, M.A., Chiu, W., and Sali, A. (2006). Refinement of protein structures by iterative comparative modeling and CryoEM density fitting. J. Mol. Biol. *357*, 1655–1668.

Topf, M., Lasker, K., Webb, B., Wolfson, H., Chiu, W., and Sali, A. (2008). Protein structure fitting and refinement guided by cryo-EM density. Structure *16*, 295–307.

Uchiumi, T., Terao, K., and Ogata, K. (1981). Identification of neighboring protein pairs cross-linked with dimethyl 3,3′-dithiobispropionimidate in rat liver 40S ribosomal subunits. J Biochem *90*, 185–193.

Ullah, H., Scappini, E.L., Moon, A.F., Williams, L.V., Armstrong, D.L., and Pedersen, L.C. (2008). Structure of a signal transduction regulator, RACK1, from Arabidopsis thaliana. Protein Sci. *17*, 1771–1780.

Valle, M., Zavialov, A., Sengupta, J., Rawat, U., Ehrenberg, M., and Frank, J. (2003). Locking and unlocking of ribosomal motions. Cell *114*, 123–134.

Vard, C., Guillot, D., Bargis, P., Lavergne, J.P., and Reboud, J.P. (1997). A specific role for the phosphorylation of mammalian acidic ribosomal protein P2. J. Biol. Chem. *272*, 20259–20262.

Warner, J.R. (1999). The economics of ribosome biosynthesis in yeast. Trends Biochem. Sci. *24*, 437–440.

Wool, I.G. (1996). Extraribosomal functions of ribosomal proteins. Trends Biochem. Sci. *21*, 164–165.

Wu, J., Beniac, D.R., and Harauz, G. (1995). Ribosomal proteins of *Thermomyces lanuginosus*—characterisation by two-dimensional gel electrophoresis and differential disassembly. Mol. Cell. Biochem. *143*, 21–34.

Wuyts, J., Van de Peer, Y., Winkelmans, T., and De Wachter, R. (2002). The European database on small subunit ribosomal RNA. Nucleic Acids Res. *30*, 183–185.

Yeh, Y.C., Traut, R.R., and Lee, J.C. (1986). Protein topography of the 40 S ribosomal subunit from *Saccharomyces cerevisiae* as shown by chemical cross-linking. J. Biol. Chem. *261*, 14148–14153.

Zuker, M. (2003). Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res. *31*, 3406–3415.

 ©2009 Elsevier Ltd All rights reserved