SUPPORTING INFORMATION (SI) APPENDIX

Structure and assembly model for the *Trypanosoma cruzi* 60S ribosomal subunit

Zheng Liu[1]*, Cristina Gutierrez-Vargas[2]*, Jia Wei[2]*, Robert A. Grassucci[1,3], Madhumitha Ramesh[4], Noel Espina[5], Ming Sun[2], Beril Tutuncuoglu[4], Susan Madison-Antenucci[5], John L. Woolford, Jr.[4], Liang Tong[2], and Joachim Frank[1,2,3#]

1. Department of Biochemistry and Molecular Biophysics, Columbia University, New York, NY 10032, USA

2. Department of Biological Sciences, Columbia University, New York, NY 10027, USA

3. Howard Hughes Medical Institute, Columbia University, New York, NY 10032, USA

4. Department of Biological Sciences, Carnegie Mellon University, Pittsburgh, PA 15213, USA

5. Parasitology Laboratory, Wadsworth Center, New York State Department of Health, Albany, NY, 12208, USA

*These authors contributed equally to this work.
#To whom correspondence should be addressed. Email: jf2192@cumc.columbia.edu

Author Contributions: J.F. and Z.L. designed the experiments. J.F. supervised the project. S.M.-A. and N.E. provided the *T. cruzi* cell extract. Z.L. and C.G.-V. purified the ribosomes. Z.L. performed the cryo-EM experiments and the image data analysis. R.A.G. assisted in the screening of the samples. J.W., Z.L., M.S., and L.T. performed the structural modeling. Z.L., C.G.-V., J.W., L.T., M.S., M.R., B.T. and J.L.W. analyzed the structure. Z.L., C.G.-V., J.W., L.T., M.S., S.M.-A., M.R., B.T., J.L.W and J.F. wrote the manuscript.

The authors declare no conflict of interest.

Data deposition: All cryo-EM map and model described in this manuscript have been deposited in the Electron Microscopy Data Bank (EMDB) (accession nos. EMD-8361, PDB code: 5T5H)

Supporting Information (SI Appendix)

SI Text: Structure and possible role of Expansion segments

Our structure shows various trypanosome-specific ESs forming contacts with functionally significant proteins that are extended or truncated in *T. cruzi*, when compared to their yeast counterparts. These ESs include ES9L, ES27L, ES7L, ES12L, and ES31L as well as those described in the main text, ES42L and the KSD.

ES7L in *T. cruzi* is composed of two long bent helices while in yeast this ES is made up of only one long helix and a short helix (Fig. S16). The tip of the trypanosome-unique long helix interacts with proteins uL30 and uL4. Protein uL30 is conserved with respect to its yeast counterpart and uL4, situated between the two long bent helices, bears a C-terminal extension (Fig. S16). Considering that uL4 is a component of the protein exit tunnel, this unusual configuration of ES7L has a possible function in sensing or regulating the passage of nascent polypeptides through the tunnel.

T. cruzi ES9L is composed of three helices instead of the two helices seen in yeast ES9L (named ES9L-h1-3). Helix ES9L-h3, with no homolog in yeast, surrounds one side of eL29, which bears an extended C-terminal portion exposed on the surface (Fig. S17). The other sides of eL29 are surrounded by ES9L-h1 and ES12. Together these three helices position the C-terminal of eL29 to be exposed stably on the surface, while the protein's N-terminal domain threads through the rRNA core in the vicinity of the PTC.

ES27L (located below srRNA1, as viewed from solvent side) and ES31L, which emerges from the base of the L1 stalk, have both increased in size compared to their yeast counterparts. The dynamic behavior of ES27L makes it visible only at low thresholds on the unsharpened density map (Fig. 1E). ES27L forms a novel intern subunit bridge as revealed in the *T. brucei* structure (1), and also provides a platform for the S-domain of the signal recognition particle (SRP) (2). Thus, it has been proposed to regulate access of factors to the peptide exit tunnel. ES31L, whose assembly is described in the main text, is anchored at the base of the L1 stalk and is a dynamic element instrumental in evacuating the exit-site tRNA (3). ES31L is composed of three helices instead of two as seen in yeast (Fig. 3B). The third, trypanosome-specific helix is in contact with protein eL8 and runs parallel with the L1 stalk. This helix also interacts with the 3'-end of the 5.8S rRNA via residue 865 (Fig. 3B). Moreover, ES31L contacts the N-terminal extension of uL23, a signaling protein for SRP (4).

Altogether, these trypanosome-specific rRNA ESs form contacts with functionally significant extended or truncated r-proteins (summarized in Table S6), suggesting the architecture of the *T. cruzi* ribosome may be specialized to modulate nascent peptide conducting, protein targeting, and other translational events such as GTPase activation by the sarcin-ricin loop.

Additional References

1. Hashem Y*, et al.* (2013) High-resolution cryo-electron microscopy structure of the Trypanosoma brucei ribosome. *Nature* 494(7437):385-389.
2. Anger AM*, et al.* (2013) Structures of the human and Drosophila 80S ribosome. *Nature* 497(7447):80-85.
3. Korostelev A, Ermolenko DN, & Noller HF (2008) Structural dynamics of the ribosome. *Current opinion in chemical biology* 12(6):674-683.
4. Park E*, et al.* (2014) Structure of the SecY channel during initiation of protein translocation. *Nature* 506(7486):102-106.

Fig. S1

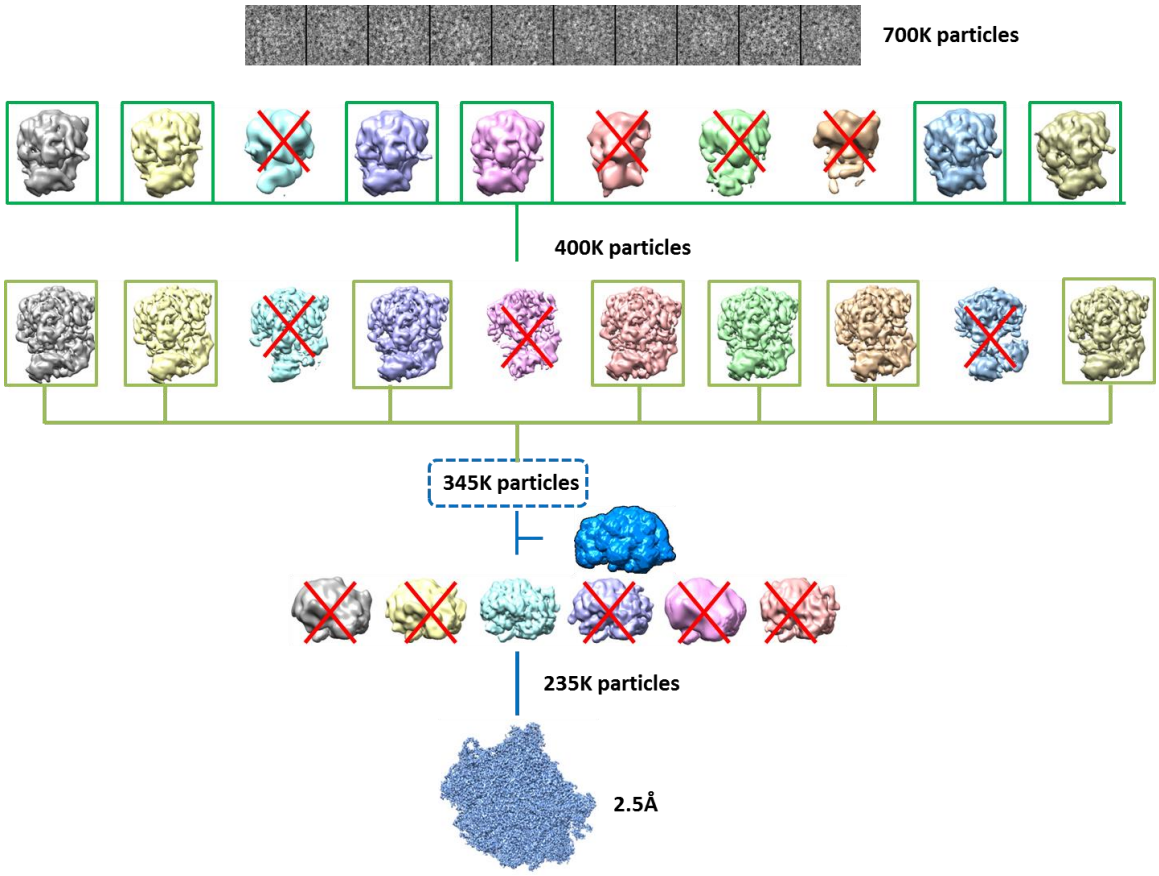

Fig. S1. Classification of the *Trypanosoma cruzi* ribosome cryo-EM data set. First row, ten particles selected from 700,000. Second row, primary classification with K= 10 classes, generating six ribosome-like and four bad classes. Third row, further classification of ribosome-like particles, again with K=10, yielding seven fully formed ribosome classes and three with broken 40S subunit. Fourth row, focused classification using K=6, with 60S subunit mask placed on all particles (345,000), yielding one dominant class. Fifth row, resulting focused reconstruction of the 60S subunit from 235,000 particles.

Fig.S2

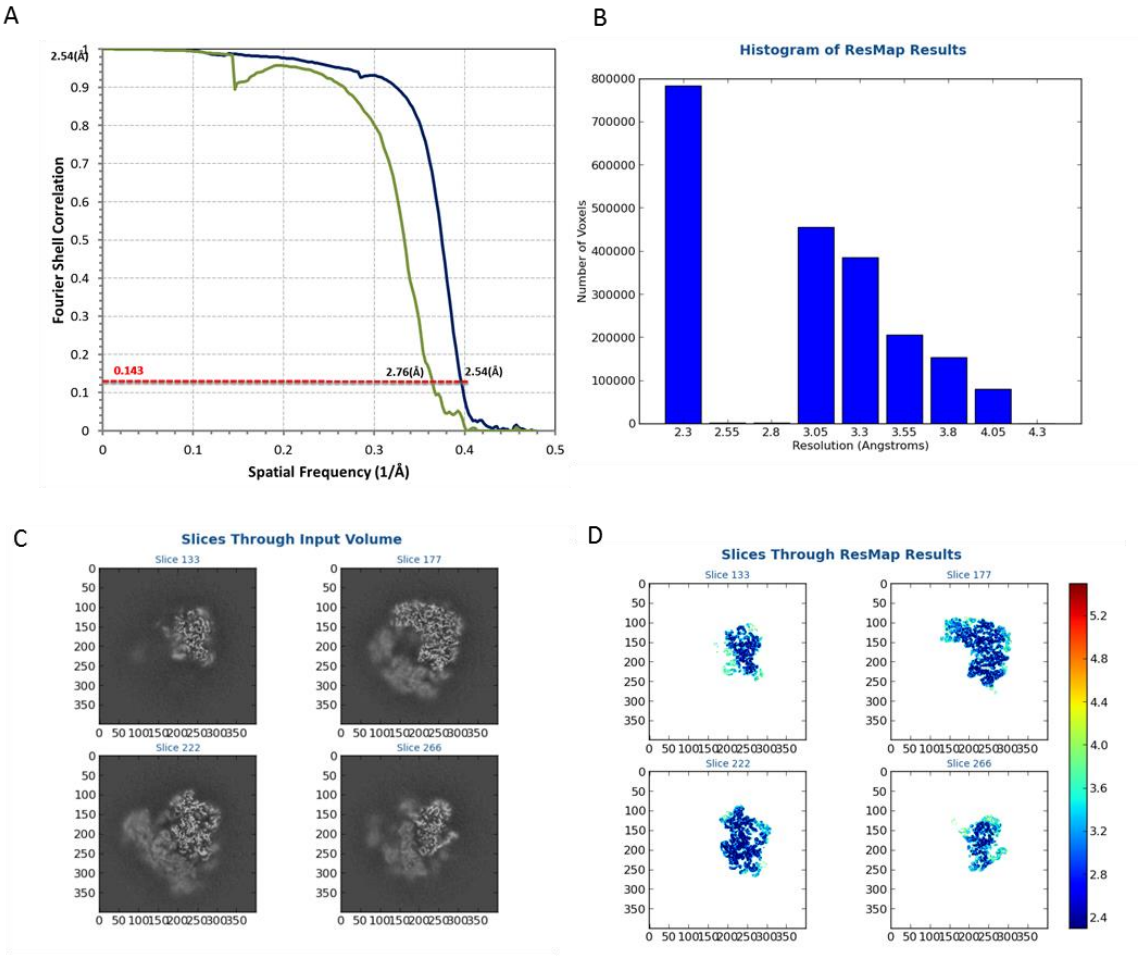

Fig S2. **Resolution estimation of the 3D-reconstruction** (A) FSC of the 60S subunit reconstruction (dark blue: model region; green: whole map). (B-D) Local resolution evaluation using *resmap*. Shown are outputs of the *resmap* software: (B) histogram distribution of resolution values; (C) density sections and (D) resolution distributions in the same set of sections shown in (C).

Fig S3

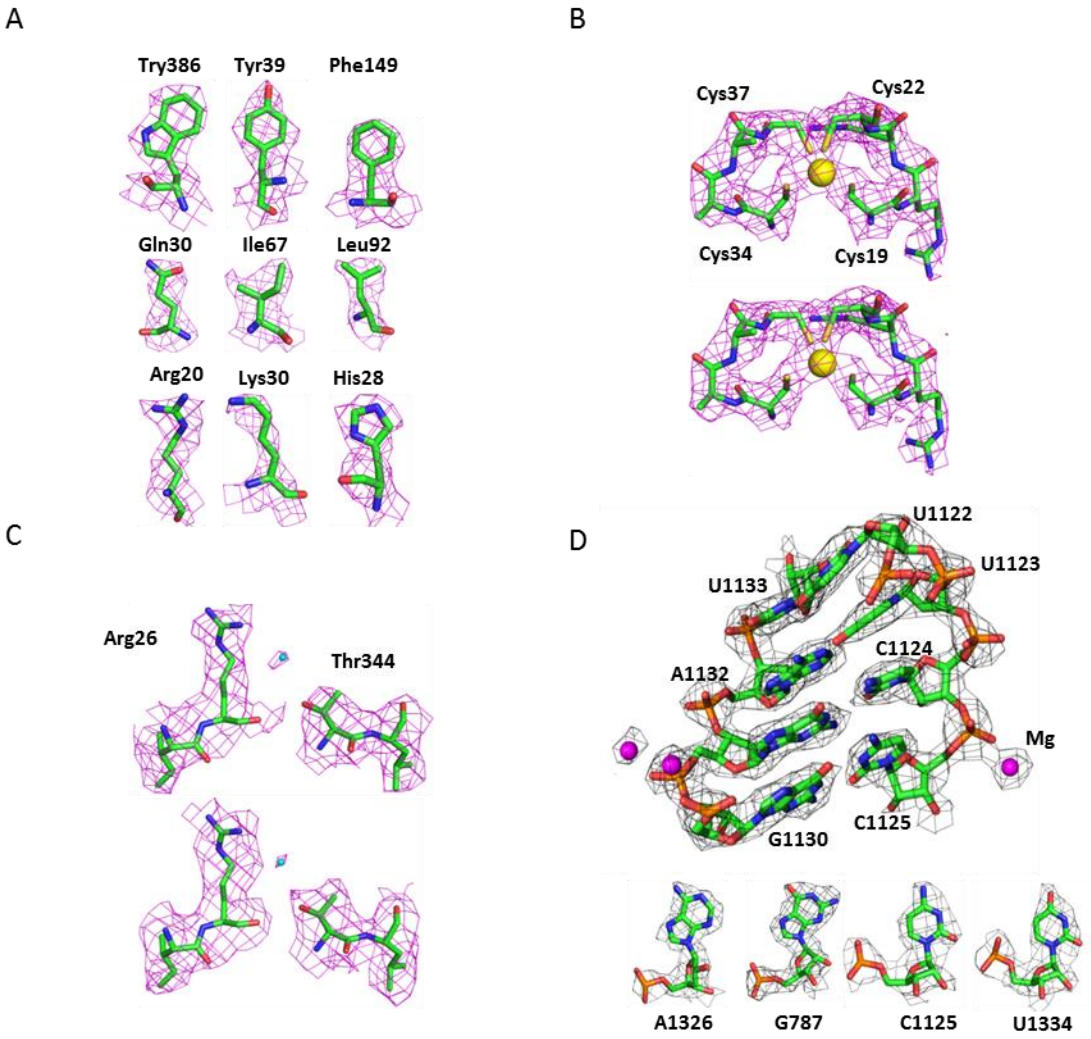

Fig. S3. Side chain features of 60S subunit of *Trypanosoma cruzi* ribosome. (A) Densities for some of the amino acids. Top, three aromatic residues; middle, three hydrophobic residues; bottom, three basic residues. These residues are from the proteins uL3, eL37 and uL15. (B,C) Reproducibility of water molecule (B), and zinc ion (C) assignments through comparison of half-maps. These regions are also displayed in Fig 1.b. (D) Cryo-EM density of a highlighted rRNA region of LSU-α docked with atomic model. Bottom panel, examples for the four nucleotides in rRNA.

Fig.S4

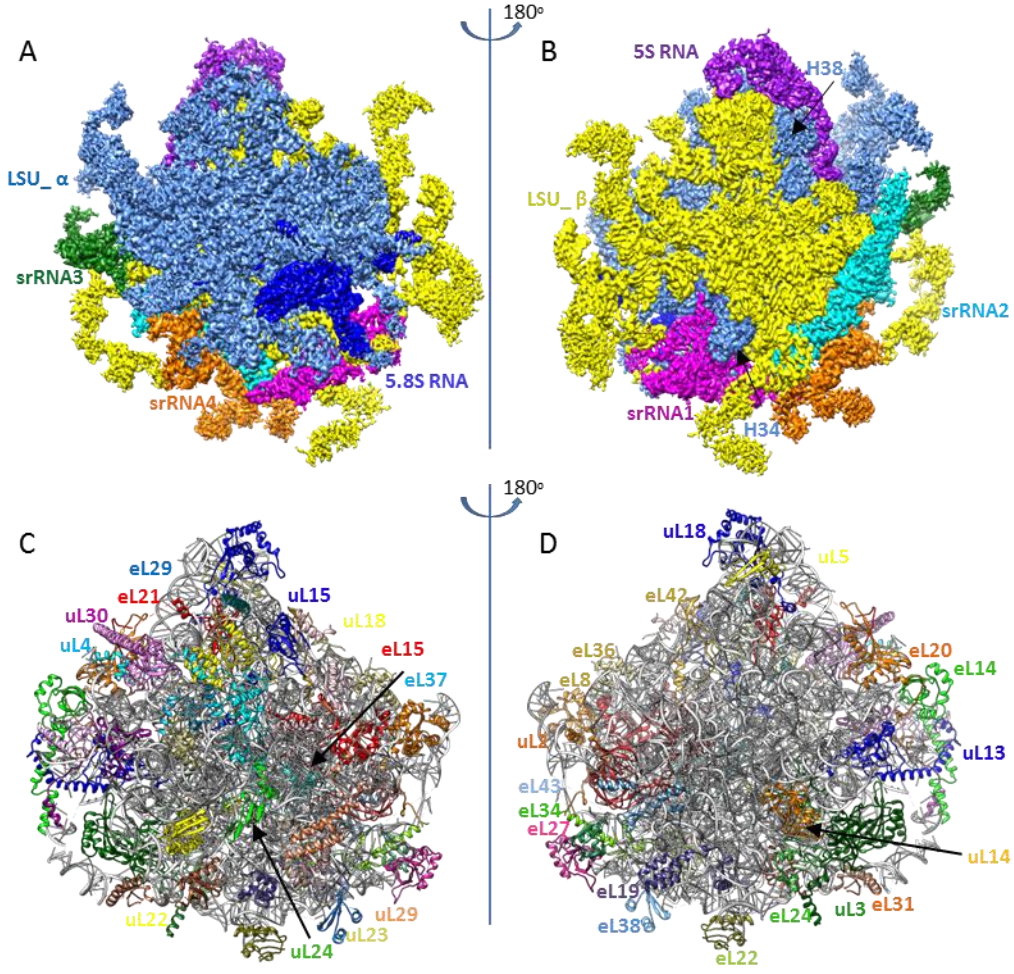

Fig. S4. rRNAs arrangement and atomic model of the 60S subunit. (A, B) Arrangement of the eight pieces of rRNA in the large subunit within the electron density map; r-proteins are omitted for clear display of rRNA. (A) Solvent and (B) interface view. (C, D) architecture of proteins in the large subunit, in solvent (C) and interface (D) view.

Fig.S5

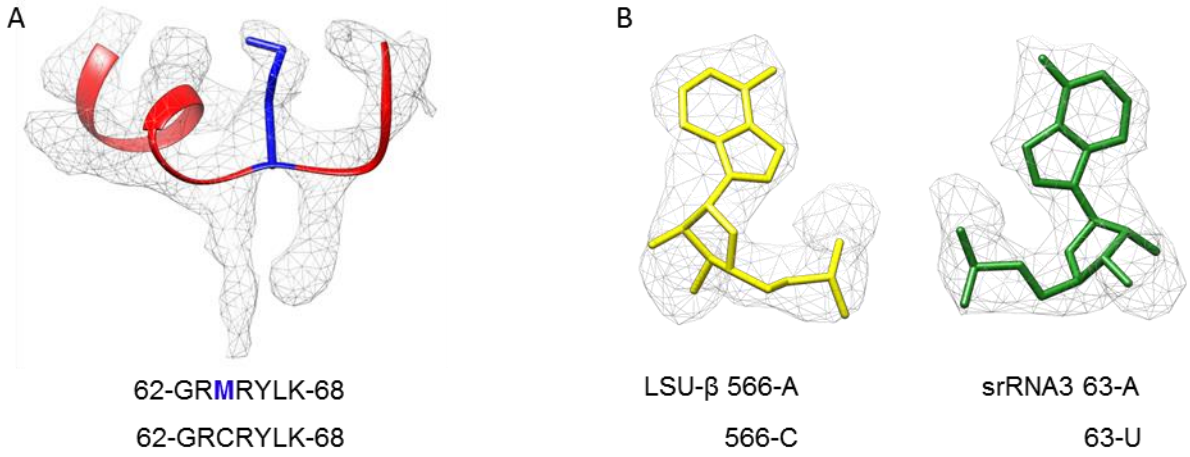

Fig. S5. Cryo-EM allows identification of variations in rRNA and proteins. (A) Density map (gray mesh) of the fragment of a.a. 62-68 on eL37. Met-64 (blue) is clearly represented in the density although the sequence XP_807547.1. has a Cysteine residue (B) Density map (gray mesh) of the A-566 on LSU-β and A-63 on srRNA3, in conflict with the sequences which contain C and U at these locations, respectively.

Fig.S6

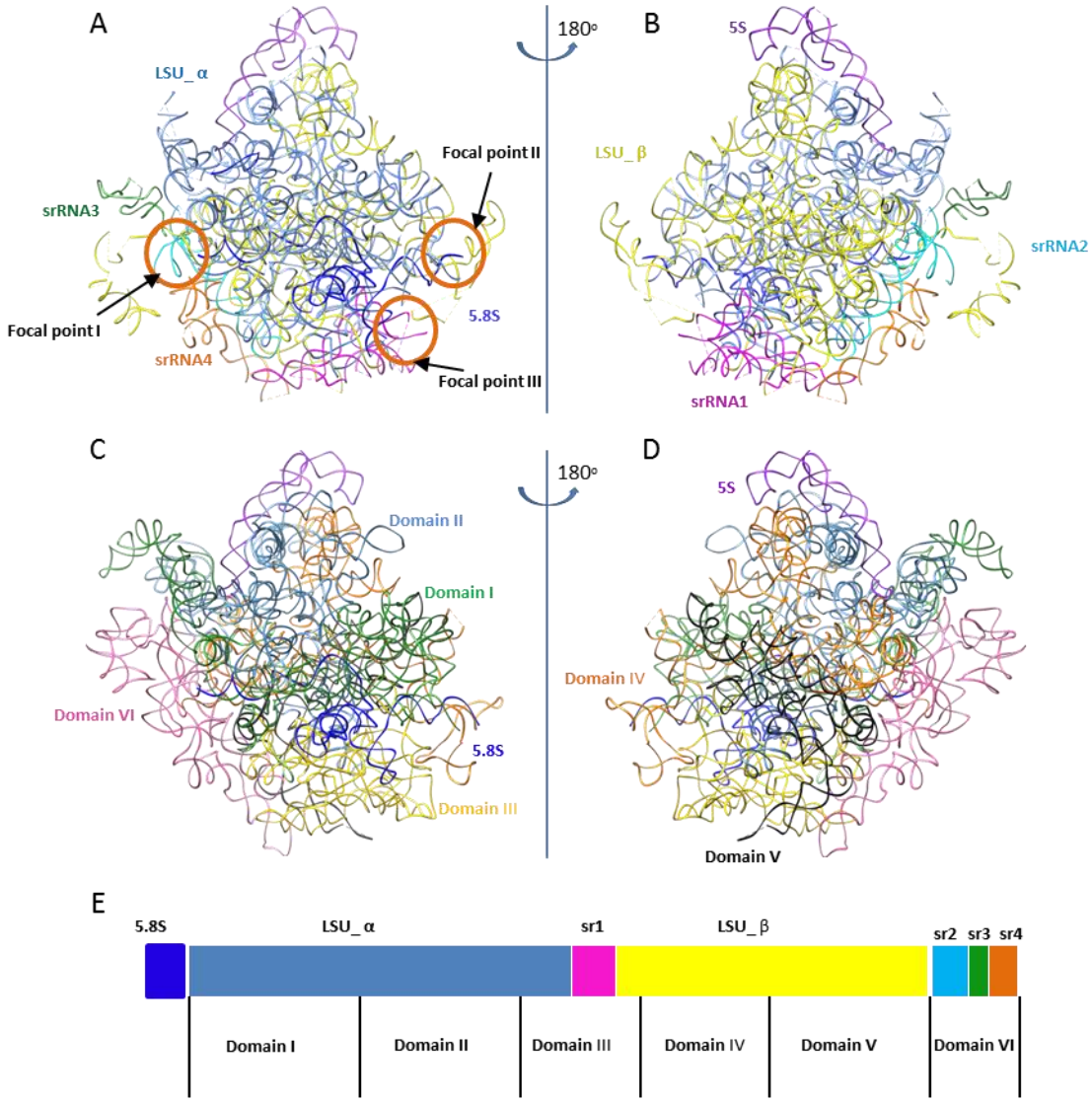

Fig. S6. *T. cruzi* large subunit rRNA pieces and the corresponding rRNA domains in yeast. (A, B) The eight pieces of rRNA in the *T. cruzi* ribosomal large subunit, in solvent (A) and interface (B) view. (C, D) The domains of rRNA in yeast 25S rRNA, in solvent (c) and interface (D) view. (E) Correspondence between the different pieces of *T. cruzi* rRNA and the domains from yeast 25S rRNA. srRNA1-4 are denoted by sr1-4.

Fig.S7

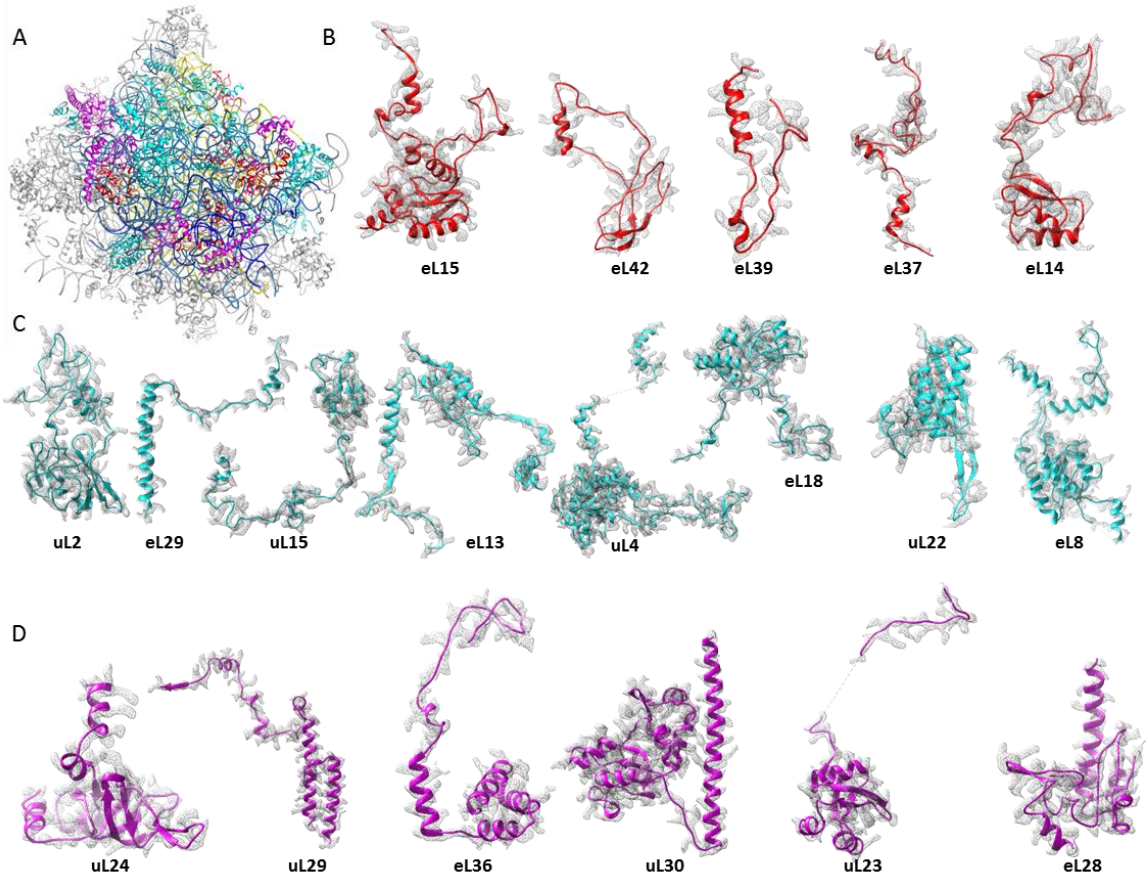

Fig. S7. Proteins for scaffold stabilization. (A) The scaffold encompasses the conserved region of LSU-α (sky blue) and LSU-B (yellow), the whole 5.8S rRNA (blue), and its associated proteins. (B) Group-one proteins, which are buried in the scaffold. (C) Group-two proteins, bearing large spanning domains. (D) Group-three proteins, located on the surface.

Fig.S8

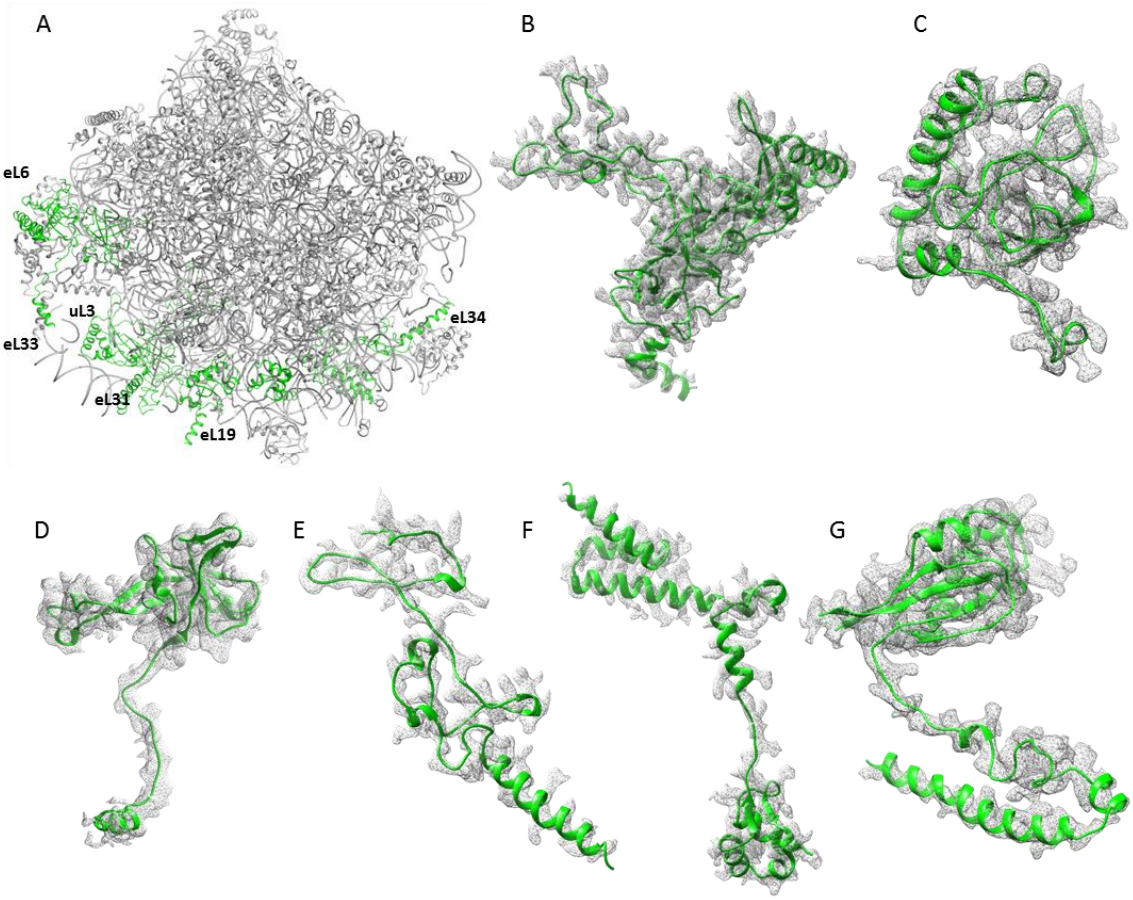

Fig. S8. Anchoring proteins for small-rRNAs assembly. (A) Locations of the anchoring proteins connecting the scaffold and small rRNAs. (B) uL3. (C) eL6. (D) eL33. (E) eL34. (F) eL19. (G) eL31.

Fig.S9

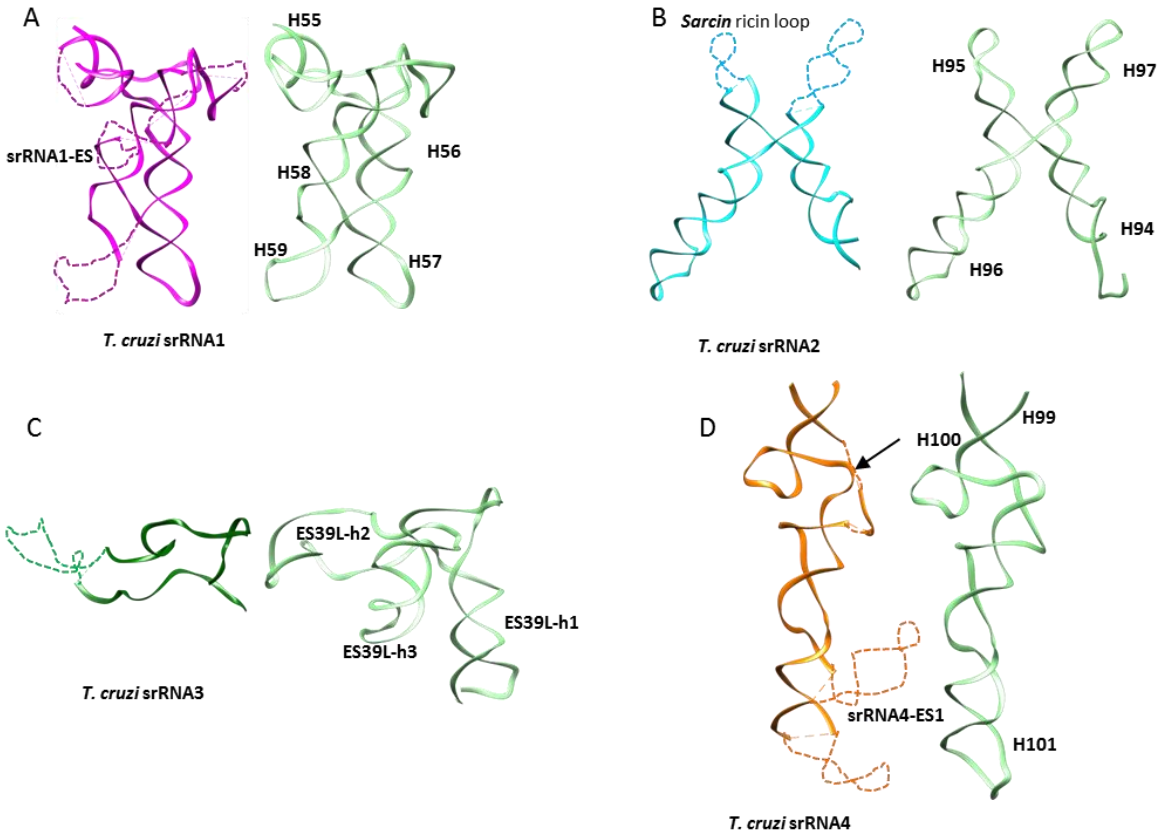

Fig. S9. Small rRNAs. (A) srRNA1 (magenta, left) and its homologous region in yeast (light green). (B) *T. cruzi* srRNA2 (cyan, left) and its homologous region, H94-97 of yeast (light green). (C), *T. cruzi* srRNA3 (forest green) and its homologous region ES39L of yeast (light green). (D) *T. cruzi* srRNA4 (orange) and its homologous region in yeast (light green). The arrow marks the shorter *T. cruzi* loop compared with yeast, which avoids conflict with the N-terminus of eL31. The un-modelled regions of the small rRNAs are connected by a dotted line according to the *T. brucei* model (4V8M.pdb) and our unsharpened density map.

Fig.S10

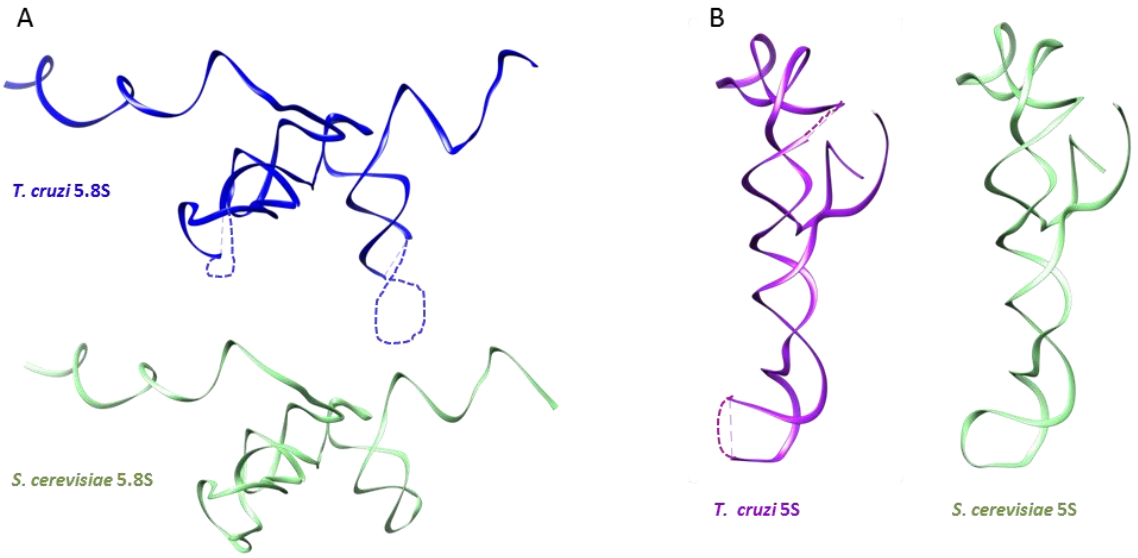

Fig. S10. *T. cruzi* and yeast 5.8S and 5S rRNAs. (A) 5.8S rRNA, upper is from our *T. cruzi* structure, bottom (light green) is from the yeast ribosome structure. (B) 5S rRNA, left is from our *T. cruzi* structure, right (light green) is from yeast. The un-modelled regions of the small rRNAs are connected by a dotted line according to the *T. brucei* model (4V8M.pdb) and our unsharpened density map.

Fig.S11

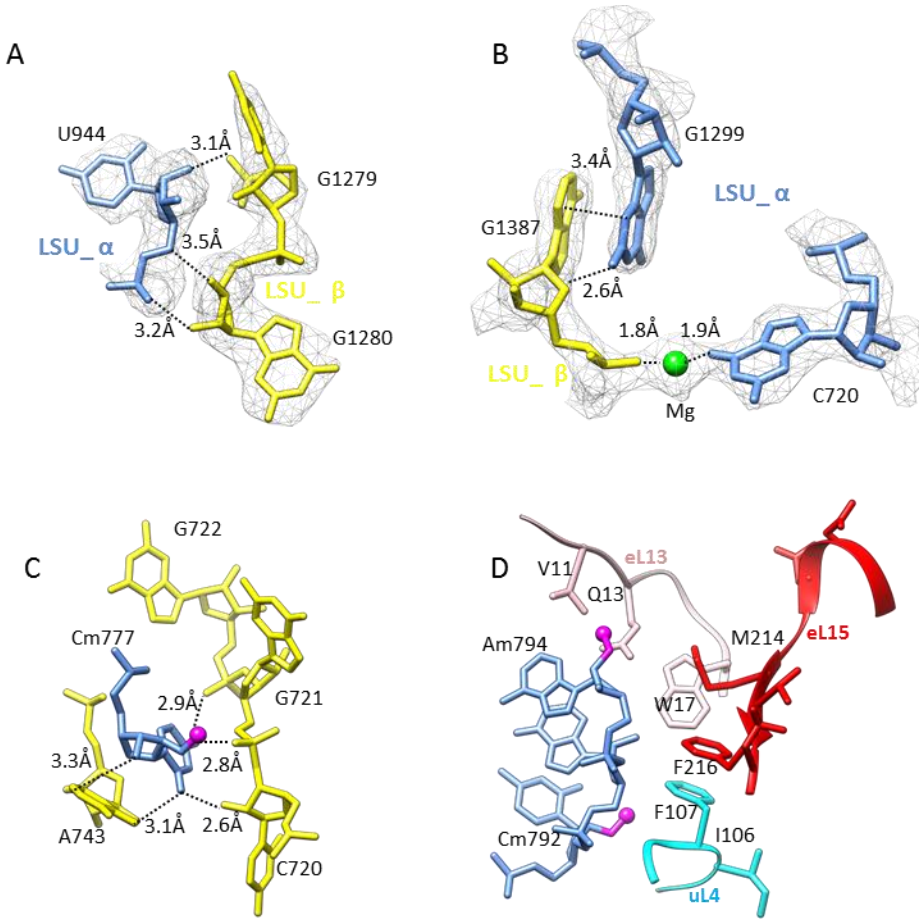

Fig. S11. rRNA and r-protein associations on the large subunit. (A) Interactions between U944 of LSU-α and C1279 and G1280 of LSU-β. (B) Interactions between A1387 of LSU-β and A1299 and G720 of LSU-α; Mg^{2+} takes part in forming the LSU-α - LSU-β interaction. (C) Modified rRNA nucleotides participate in forming interactions between LSU-α and LSU-β. Cm777 is a methyl-modified residue of LSU-α. (D) A hydrophobic region formed by modified nucleotides and r-proteins. The magenta spheres in rRNA model denote methyl groups.

Fig.S12

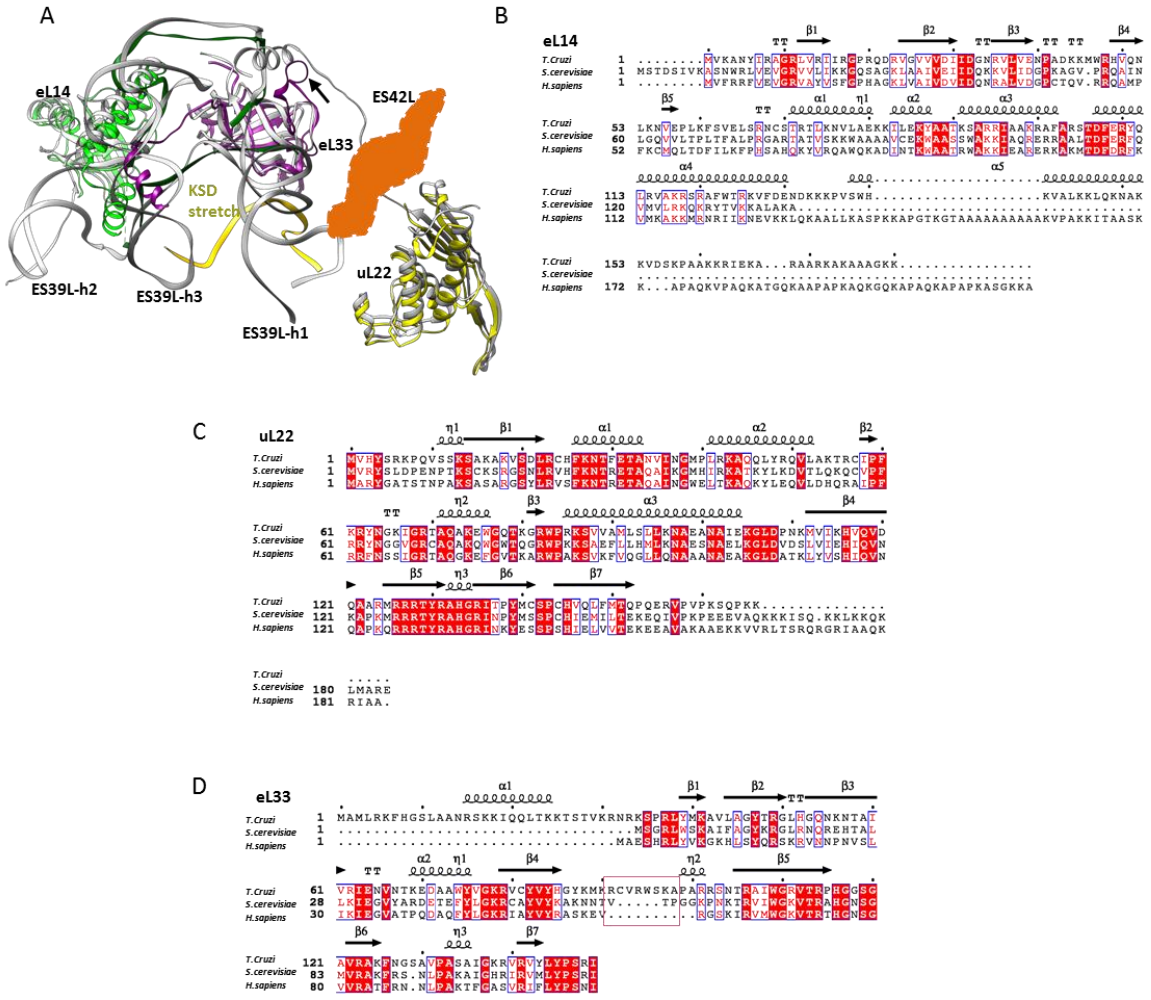

Fig. S12. srRNA3 and its homologous rRNA ES39L in yeast surroundings. (A), Overlay of srRNA3 and its surrounding region, which includes eL14, eL33, and uL22 of *T. cruzi*. For comparison, the ES39L region of yeast is also shown in grey. ES42L has been added by manual drawing. The arrows indicate the insertion of the eL33 (marked by magenta box in (D)), which occupies the position of the C-terminal of uL22 in the yeast ribosome structure. (B) Sequence alignment of r-protein eL14. (C) Sequence alignment of r-protein uL22. (D) Sequence alignment of r-protein eL33.

Fig. S13

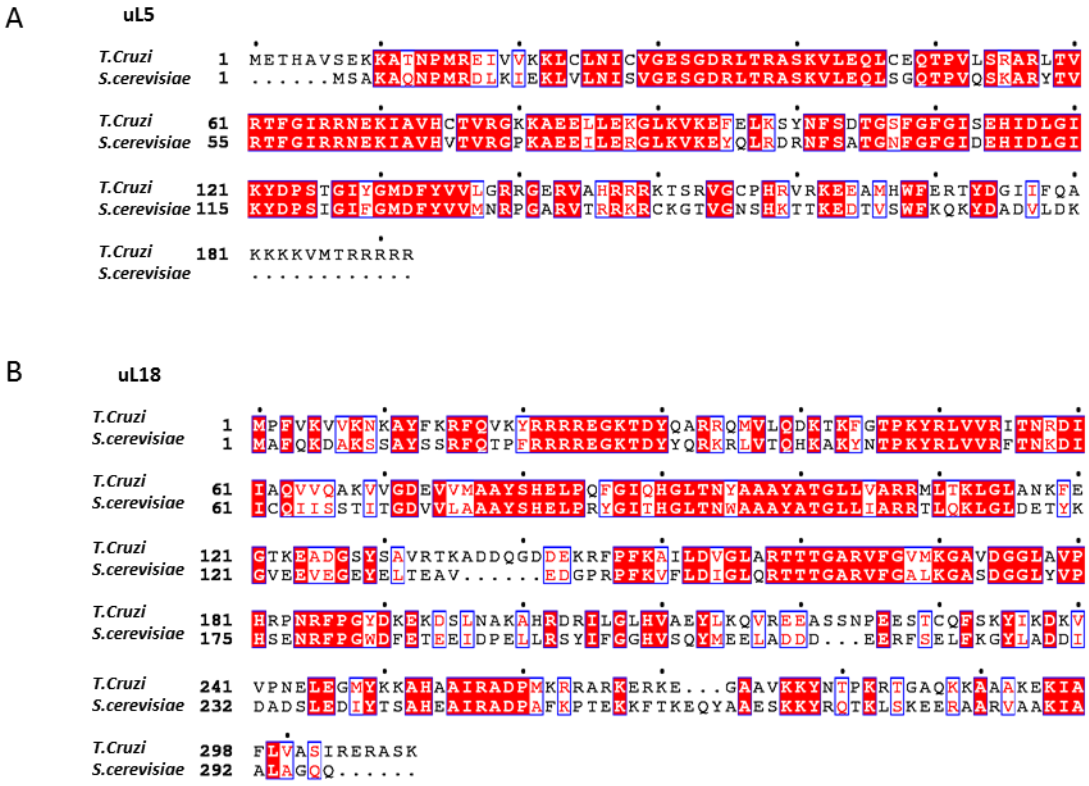

Fig. S13. Sequence alignment of 5S rRNA-associated proteins. (A) Sequence alignment of protein uL5, whose identity between the two sequences is 69%. (B) Sequence alignment of protein uL18, whose identity between the two sequences is 49%.

Fig. S14

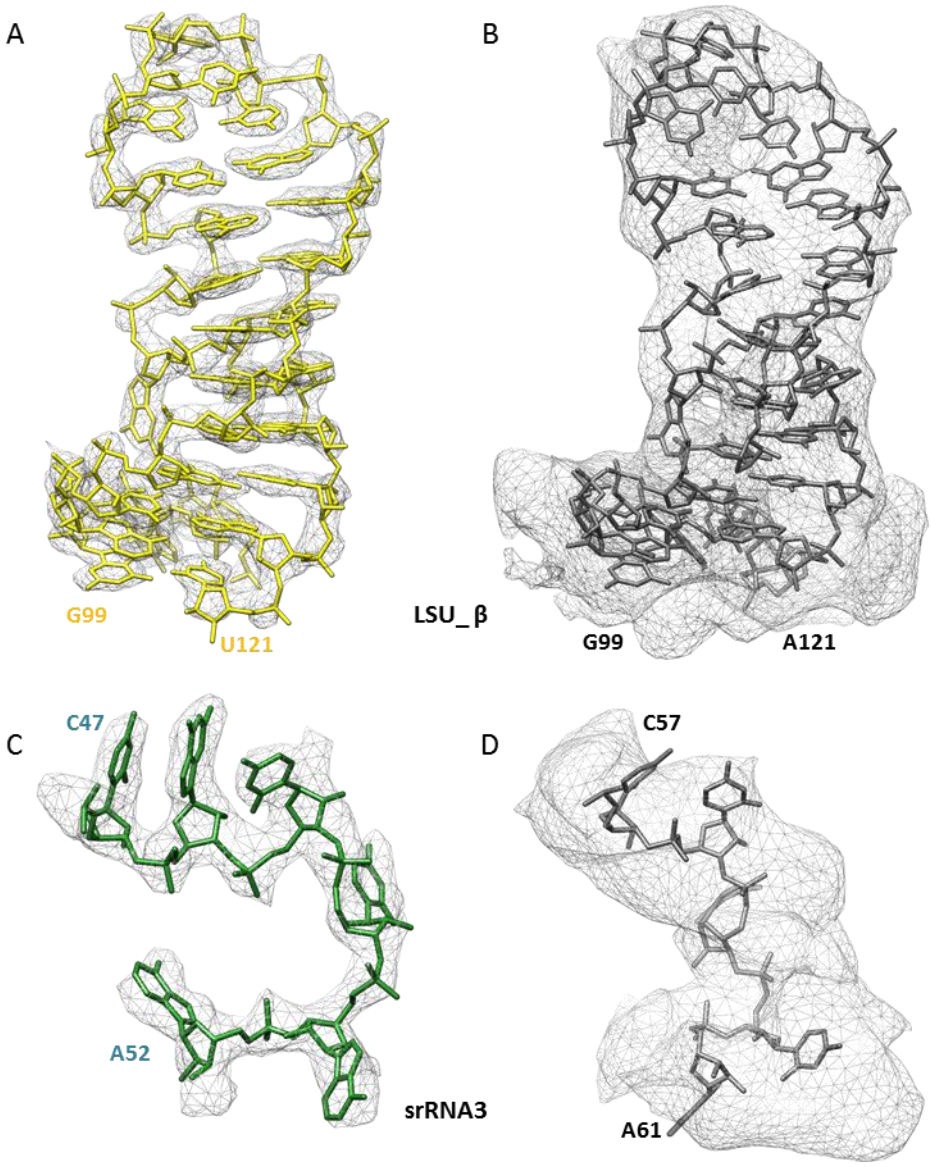

Fig. S14. Comparison of some regions of 2.5 Å *T. cruzi* and 5.5 Å *T. brucei* ribosome density maps, superimposed with their corresponding atomic models. (A) A double-helical region of LSU-β from the *T. cruzi* map and model. (B) The same region from the *T. brucei* map and model as in (A). (C) A single-stranded fragment of srRNA3 from the *T. cruzi* map and model. (D) The same region from the *T. brucei* map and model as in (C).

Fig. S15

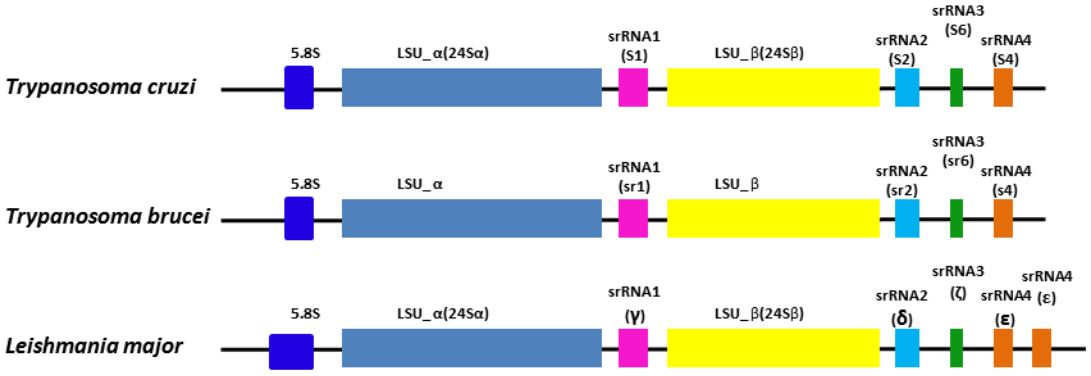

Fig. S15. Schematic representation of the large subunit rRNA genes (except 5S rRNA) from trypanosomatids. The names in the bracket are their respective old nomenclatures.

Fig. S16

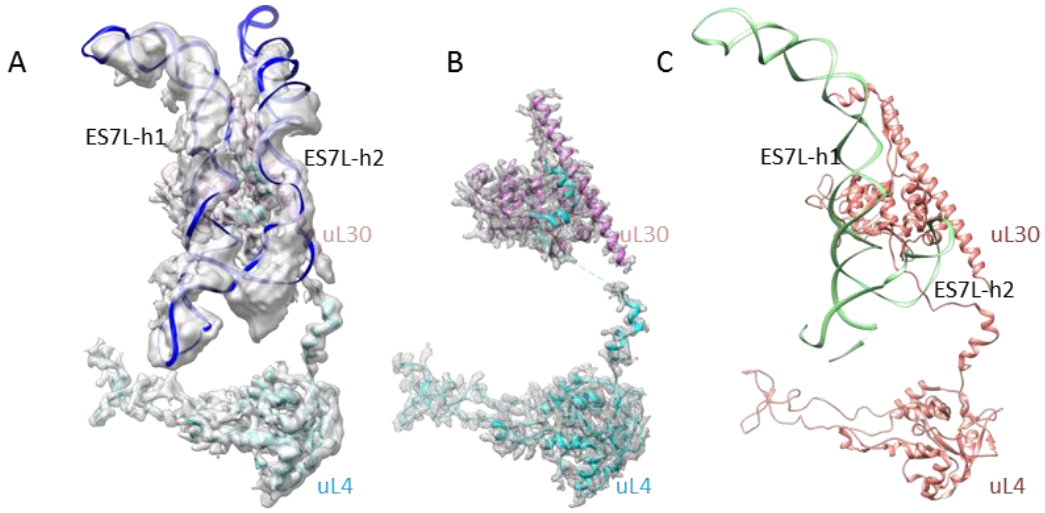

Fig. S16. ES7L and its contacting proteins uL4 and uL30. (A) ES7L (blue) taken from the *T. brucei* model (4V8M.pdb), and uL4 and uL30 from our model were fitted to the *T. cruzi* unsharpened map. (B) uL4 and uL30 superimposed onto the sharpened map, in the same view as in (A). The extension of uL4 is gripped by the long helixes of ES7L, as shown by combination of (A) and (B). (C) ES7L (light green), composed of one long bent helix and one short helix, and uL4 and uL30 (brown) from yeast shown in the same view as (A) and (B).

Fig. S17

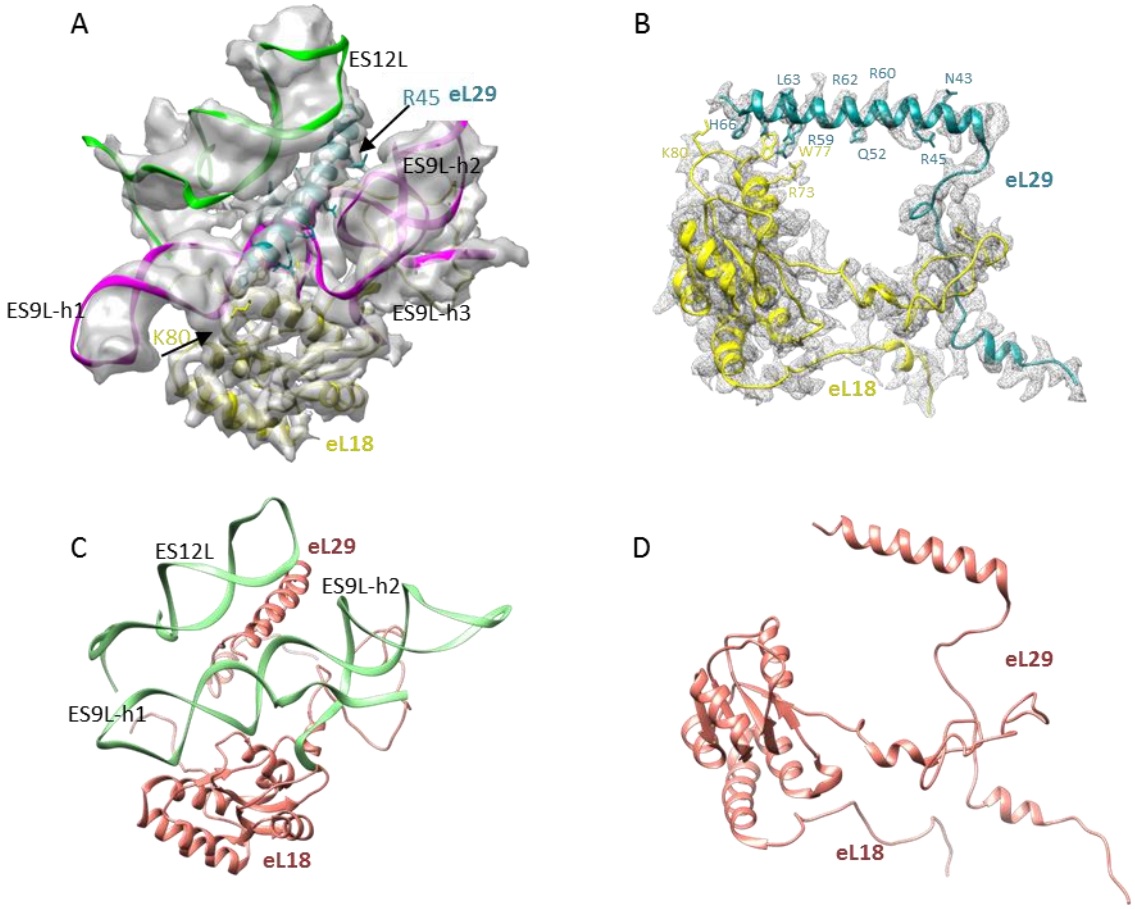

Fig. S17. C-terminal extension of eL29 contacts eL18 and is surrounded by ES9L and ES12L. (A) ES9L (magenta) and ES12L (green) taken from the *T. brucei* ribosome model (4V8M.pdb), and eL18 and eL29 from our model were fitted into the *T. cruzi* unsharpened map. (B) Interactions of eL29 extension with eL18. The models are fitted to the sharpened map. (C) ES9L, ES12L, eL18, and eL29 in yeast, in the same view as in (A). (D) eL29 and eL18 of yeast in the same view as in (B).

Table. S1

	5.8S	LSU- α	srRNA1	LSU- β	srRNA2	srRNA3	srRNA4	Ref.
T. cruzi	172	1969	209	1661	178	72	136	[11]
T. brucei	115	1840	220	1570	180	70	140	[8]
L. major	283	1781	213	1525	183	73	129	[5]

Table S1. The length of different pieces of rRNA (except 5S rRNA) for three trypanosomatids

Table. S2

Parameters	Value
3D reconstruction	
Particle mumbles	235 348
accuracy of rotation (˚)	0.43
accuracy of translation (Å)	0.42
Resolution (unmasked, Å)	2.9
Resolution (masked, Å)	2.5
Map sharpening B-factor (Å^2)	-51.16
Model Refinement	
Cell dimensions a=b=c (Å)	240.35
Cell dimensions α =β=γ (˚)	90
Map CC	0.58
Resoluion(Å)	2.5
r.m.s. deviations Bond length (Å)	0.016
r.m.s. deviations Bond angle (˚)	1.5
Clash score	18.8
Ramachandran plot	
Favored (%)	93.4
Allowed (%)	5.4
Outliers (%)	1.2
MolProbity Score	2.2
Favored rotatomers (%)	98.0
Corrected sugar puckers (%)	97.6
Good backbone conformation (%)	75.2
EM Ringer Score	4.32

Table S2. 3D reconstruction and model statistics of *T. cruzi* large ribosomal subunit

Table S3

Components	Number
Proteins	38
rRNA	8 pieces
H20	84
Mg	105
Zn	3
Methylation	66 2'-O methylation
	10 nucleobase methylation

Table S3. Summary of the *T. cruzi* ribosomal large subunit model

Table S4

RNA	Length	Built
5S	118	1-87,91-118
5.8S	172	1-79,89-125,139-169
LSU-α	1969	3-154,175-198,205-219,232-281,304-519,585-604,744-853,866-880,916-1171,1221-1384,1491-1545,1569-1691,1709-1805,1918-1964
srRNA1	209	2-12,21-74,95-152,175-185,191-209
LSU-β	1661	6-136,446-594,616-789,852-913,1105-1113,1134-1232,1240-1330,1343-1397,1443-1580, KSD (partially).
srRNA2	178	21-44,57-128,154-178
srRNA3	72	1-17,44-72
srRNA4	136	2-47,62-68,87-112,120-125,128-133

Table S4. Summary of the modeled ribosomal rRNA

Table S5

LSU-α	Location	LSU-β	Location
154-175	H16	**137-445**	ES27L
198-232	ES42L	**595-615**	H69
281-304	H17	**790-851**	H76
519-584	ES7L	**914-1104,1114-1133**	ES31L
853-866,880-916	ES9L	**1233-1249**	H84
1171-1221	H38	**1331-1342**	H87
1385-1491	H42-43	**1398-1442**	H89
1545-1569	ES15L		
1691-1709	*T.c.* specific fragment		
1805-1917	ES19L		

Table S5. The unmolded ribosomal rRNA of LSU-α /β location

Table S6

ES	Coordinates	Function	Uniqueness compared with yeast ribosome	Description
ES3L	5.8S (121-139)	Structural stabilization	Increased size	Contacts the 3' end of LSU-α
ES7L	LSU-α (484-750)	Regulation	two long bent helices	Surrounds uL4 linking ES7L to PET
ES9L	LSU-α(848-953)	Regulation	One additional helix	Surrounds eL29 linking ES9L to PTC
ES27L	LSU-β(128-451)	Structure stabilization and regulation	Longer	Contacts S-domain of SRP and forms a bridge with SSU
ES31L	LSU-β(866-1146)	Structure stabilization and regulation	One additional helix	Contacts the 3'end of 5.8S and uL23, which contacts SRP
ES42L	LSU-α(198-239)	Structure stabilization and regulation	Trypanosome specific Occupies the C-termini of yeast uL22	Contacts srRNA2-4 cleavage sites and contacts uL22(part of PET)
KSD	LSU-β(1585-1661)	Regulation	Trypanosome specific	Binds to uL3 which "wiggle" to PTC (19) and contacts srRNA2 containing sasin-ricin loop

Table S6. The expansion segments with unusual size and Trypanosome-specific ES

Table S7

Protein names	Old nomenclature	Protein ID	Range modeled	a.a. size	Modeled percentage	ChainID
uL2	L2	XP_816366.1	2-246	260	94%	e
uL3	L3	XP_814673.1	2-404	428	94%	f
uL4	L4	XP_804516.1	6-302,315-342	374	87%	r
uL5	L11	XP_803249.1	17-27,73-78,105-109,130-172	192	34%	L
eL6	L6	XP_806123.1	23-104,144-193	193	68%	v
eL8	L7a	XP_821284.1	88-169,185-310	319	65%	x
uL13	L13a	XP_810252.1	20-222	222	91%	O
eL13	L13	XP_810966.1	2-134,168-206	218	79%	N
uL14	L23	XP_821640.1	13-139	139	91%	W
eL14	L14	XP_804656.1	5-160	180	87%	P
uL15	L27a	XP_821642.1	2-145	145	99%	b
eL15	L15	XP_814692.1	2-204	204	100%	Q
uL18	L5	XP_805667.1	6-113, 146-259	309	62%	u
eL18	L18	XP_819826.1	2-193	193	99%	I
eL19	L19	XP_820996.1	3-189	372	50%	T
eL20	L18a	XP_819631.1	2-178	179	99%	S
eL21	L21e	XP_812750.1	3-113,134-148	159	79%	U
uL22	L17	XP_804874.1	3-154	166	92%	R
eL22	L22	XP_819834.1	24-33,46-72,75-122	130	65%	V
uL23	L23a	XP_820993.1	78-94,102-193	194	56%	X
uL24	L26	XP_806321.1	7-119	143	79%	Z
eL24	L24	XP_804713.1	3-63	125	49%	Y
eL27	L27	XP_809804.1	2-50,58-86,107-133	133	79%	a
eL28	L28	XP_820211.1	5-82,88-128	146	82%	c
uL29	L35	XP_811059.1	7-38,46-126	127	89%	k
eL29	L29	XP_803994.1	6-68.	71	89%	d
uL30	L7	XP_816099.1	28-242	242	89%	w
eL30	L30	XP_810701.1	14-72,91-97	105	63%	g
eL31	L31	XP_808028.1	13-156,164-186	188	89%	h
eL32	L32	XP_817439.1	11-123	132	86%	i
eL33	L35a	XP_821812.1	8-149	149	95%	l
eL34	L34	XP_811573.1	4-107	171	61%	j
eL36	L36	XP_805417.1	14-108	114	83%	m
eL37	L37	XP_807547.1	2-82	84	96%	n
eL38	L38	XP_816521.1	2-14,21-59	82	63%	p
eL39	L39	XP_808953.1	2-51	51	98%	q
eL42	L44	XP_806035.1	2-94	106	88%	t
eL43	L37a	XP_808231.1	3-87	90	94%	o

Table S7. Modeled proteins of the *T. cruzi* 60S ribosomal subunit

Table S8

srRNAs	Anchoring proteins	Contacting sites
srRNA1	eL19	92,103,110,124,128
	eL34	W47, H51
srRNA2	uL3	26,27,31,50,92,99,104,161,130,131,133,185,229,236,339,
srRNA3	eL6	47,73,81,76-78,170
	eL33	85,87,139,39,141,143,109,79,79,145,31-33
srRNA4	eL31	91,115,146

Table S8. Residues of anchoring proteins involved in assembly of small pieces rRNA

Table S9

Expansion Segment	Comparison between *T. cruzi* 60S and human ribosome
ES3L	A couple of nucleotides longer in *T. cruzi*
ES4L	Conserved
ES5L	A couple of nucleotides shorter in *T. cruzi*
ES7L	Truncated in *T. cruzi*
ES8L	A couple of nucleotides shorter in *T. cruzi*
ES9L	Conserved, both composed of three helices but have different configurations
ES10L	Truncated in *T. cruzi*
ES12L	Conserved
ES15L	Truncated in *T. cruzi*
ES19L	Expanded in *T. cruzi*
ES20L	Missing in trypanosomes/location at cleavage site LSU-α and srRNA1
ES26L	Missing in trypanosomes/location at cleavage site srRNA1 and LSU-β
ES27L	Truncated in *T. cruzi*
ES30L	Truncated in *T. cruzi*
ES31L	Expanded in *T. cruzi*
ES35L	Conserved
ES39L	Truncated in *T. cruzi*
ES41L	Expanded in *T. cruzi*
ES42L	Trypanosome specific

Table S9. Expansion segments compared with human ribosome counterparts

Table S10

Protein names	a.a. size of *T. cruzi*	a.a. of *S. cerevisiae*	a.a. size of *H. sapiens*	Description
uL2	260	254	257	Very conserved
uL3	428	387	403	Insertion of a.a 209-211 and extension on C-terminal contacts sRNA4
uL4	374	362	427	C-terminal Extension
uL5	192	174	178	N- and C- terminal extensions
eL6	193	176	192	Inserted residues 128-143 are located on the less ordered region.
eL8	319	256	288	N-terminal extension
uL13	222	199	203	Contacts srRNA3
eL13	218	199	211	Conserved
uL14	139	137	140	Conserved
eL14	180	138	215	Shortened at N-terminal and C-terminal extension compared to S.c
uL15	145	149	148	Conserved
eL15	204	204	204	Conserved
uL18	309	297	297	Insertion a.a 136-141, a.a 222-224, deletion of residues 163-165 in yeast, and C-terminal extension
eL18	193	186	188	Insertion of a.a 72-77
eL19	372	189	196	Conserved, C-terminal extension contacts SSU
eL20	179	174	176	Conserved
eL21	159	160	160	Conserved
uL22	166	184	184	C-terminal is shortened, resulting in absence of contacts with ES39L homolog, srRNA3
eL22	130	121	128	Conserved
uL23	194	142	156	Extension at N-terminal
uL24	143	127	145	Extension at N-terminal
eL24	125	155	145	Shortened at C-terminal
eL27	133	136	136	Conserved
eL28	146	/	137	Insertion of a.a 118-132
uL29	127	120	123	Insertion of a.a 81--84
eL29	71	59	159	Extension at C-terminal compared to yeast
uL30	242	244	248	Conserved
eL30	105	105	115	Conserved
eL31	188	113	125	Extension at C-terminal. Contacts KSD
eL32	132	130	135	Conserved
eL33	149	107	110	Extension at N-terminal and insertion of a.a 89-98
eL34	171	121	117	Extension at C-terminal and insertion of a.a 45-50 that contact srRNA1
eL36	114	100	105	Conserved
eL37	84	88	97	Conserved
eL38	82	78	70	Contacts srRNA1
eL39	51	51	51	Conserved
eL42	106	106	106	Conserved
eL43	90	92	92	Conserved

Table S10. Proteins compared with human and yeast ribosome counterparts