

Supplementary Online Material

Cryo-EM Structure of the Archaeal 50S Ribosomal Subunit in Complex with Initiation Factor 6 and Implications for Ribosome Evolution

Basil J. Greber, Daniel Boehringer, Vlatka Godinic-Mikulcic, Ana Crnkovic, Michael Ibba, Ivana Weygand-Durasevic, Nenad Ban^{*}

^{*} To whom correspondence should be addressed. E-mail: ban@mol.biol.ethz.ch, phone: +41 44 633 27 85, fax: +41 44 633 12 46.

Supplementary Text

Analysis of Conserved Residues in IF6

The most highly conserved residues in IF6 do not cluster on the ribosome binding surface of IF6 and are mostly buried (Fig. S3b), suggesting that most of them are required for IF6 stability and folding rather than making ribosome binding. Based on genetic studies in yeast, the mutation of several of these conserved residues has been reported to reduce the stability of eIF6 binding to the large ribosomal subunit¹. This includes the residues corresponding to the universally conserved Gly14, Gly105 and Asn106 (*S. cerevisiae* numbering¹). The glycines are located on the inward facing side of the ring of helical segments that forms part of the ribosome interacting surface of IF6 (Fig. S3b). Mutations in these residues may structurally distort IF6 in a way that reduces the interactions of nearby residues with the ribosomal subunit. Asn106, in contrast, is located on the surface of IF6 and contacts the ribosome. In addition to these highly conserved residues, several less strictly conserved residues found on the ribosome binding face of IF6 are also affected by mutations that weaken the interaction between the ribosomal subunit and IF6¹.

Possible Causes for the Loss of Ribosomal Proteins in Euryarchaeota

In archaea, the distribution of ribosomal proteins exhibits a domain wide trend towards a more and more reduced ribosomal proteome in late branching *Euryarchaeota*². Some cases of ribosomal protein loss in bacteria and eukaryotes have been observed in organisms that have undergone genome reduction². Genome reduction is mostly associated with specific lifestyles of the affected organisms as obligate pathogens^{3; 4; 5; 6}, endosymbionts⁷, or obligate parasites⁸, but has also been reported for free living mesophilic bacteria^{9; 10}. However, the reduction of the ribosomal proteome in *Euryarchaeota* does not seem to correlate with genome size (preliminary analysis, data not shown), suggesting that the reduction of the ribosomal proteome is not caused by genome compaction. This does not exclude the possibility that genome compaction is responsible for the loss of ribosomal proteins in some species, but suggests that it may not be the cause of the domain-wide trend² of ribosomal protein loss in archaea.

Possibly, ribosome assembly or function is influenced by the physico-chemical intracellular conditions and growth regimes associated with the extreme lifestyles of many archaea, contributing to the observed decline in the number of ribosomal proteins throughout euryarchaeota. Some ribosomal proteins might no longer be required in some environments, or evolution of a more compact, more stable ribosomal particle might confer some selective advantages. Additionally, the constraints imposed on the amino acid composition and structure of proteins in extreme environments, such as acidification of the proteome in halophiles, may be a contributing factor^{11; 12; 13}.

Supplementary Figures

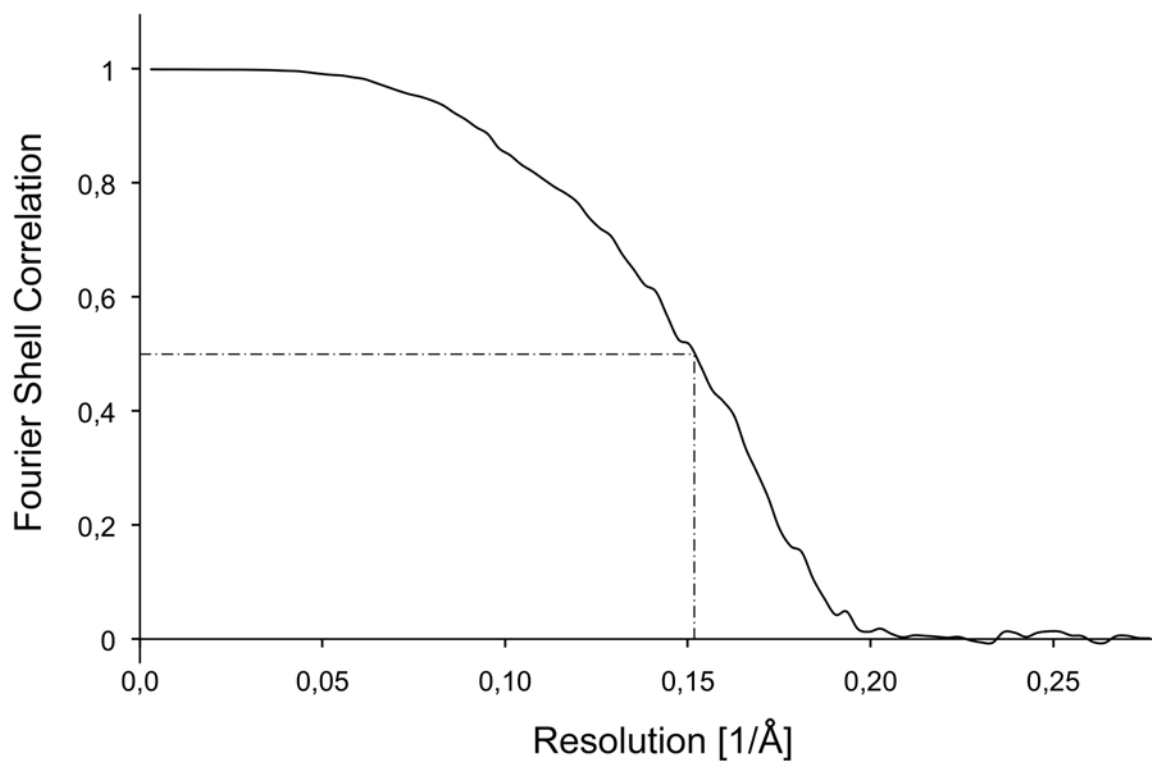


Fig. S1. Fourier Shell Correlation (FSC) curve of the final contoured 50S-aIF6 cryo-EM map calculated from masked density maps. The resolution estimate is 6.6 Å at FSC = 0.5 and 5.5 Å at FSC = 0.143¹⁴. Based on the features observed in the cryo-EM map, the FSC = 0.5 criterion is applied for its interpretation throughout the manuscript.

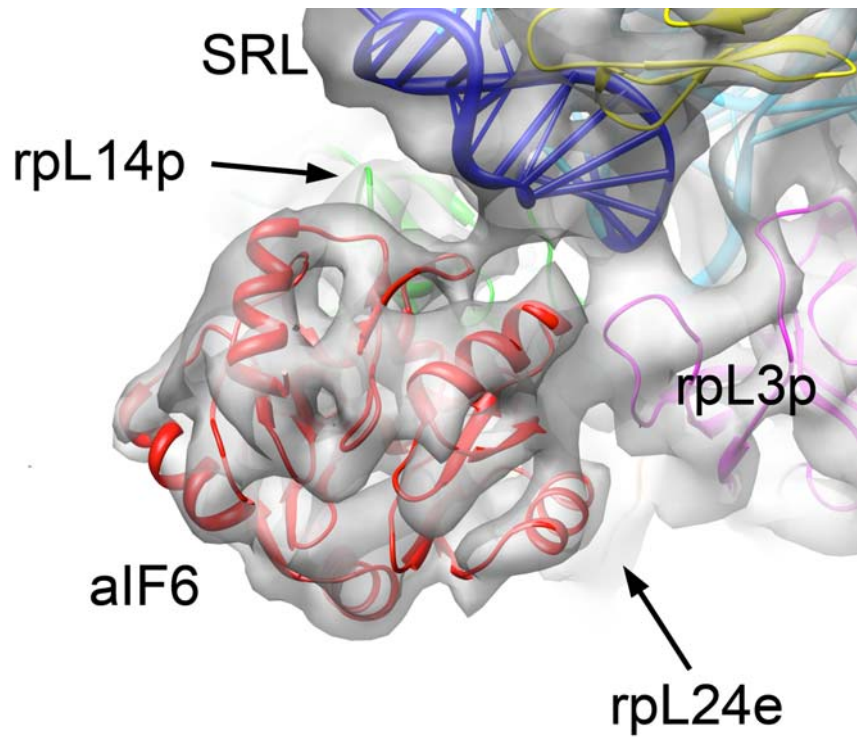


Fig. S2. Additional view of the binding site of aIF6 on the 50S ribosomal subunit (rotated from Fig. 3c). The Sarcin-Ricin-Loop is depicted in deep blue, aIF6 in red, rpL14p in green, rpL3p in magenta, and rpL24e in orange.

(a)

```
M. thermautotrophicus 1 --MIRRLINLSGNPNLGVYISVTDVSVALPQNTPEKFEGLVREAA--EVEVLKVSISGSSINGALAVGSSNGF 68
M. jannaschii 1 MTMIIRKYFSGIPTGVLLATTEEITLLPIFLDKDDVNEVSEVL--ETKCLQNTIGGSSLVGSLSVANKYGL 70
A. fulgidus 1 ---MKVAVNGNPLIGLYAKVSEYAYVVG--VNHLEPLDAIQEKL--DVDVIVTKIAGSELVGGAMMALSRGA 65
P. horikoshii 1 --MHIERLDFENSPYLGVFVATDRVLLIREGLQEKLEVIRETL--KVPVIEVSVKSRITGLTATCSNAI 69
S. acidocaldarius 1 --MILDKLSIFGTDNIGIYIFTNDKTYTIPKIDDKVEIEKIQGILL--KTEIQTTSKSLVGLVLTGNDVI 69
D. discoideum 1 --MATRLQYENSCDVGVLKLTNKYCLVQCGSKQFLHTVENRDLADHPIPVVETSAGTRIVGRLSAGNKNGL 70
Z. mays 1 --MATRIQFENNCEIGVFSKLTNAYCLVAIGGSENFYSAFAEALADAPIPVVKTSGGTRIGRLCVGNKNGL 70
O. lucimarinus 1 --MAARHQFENSNEIGVFAALTNAYCLAAGVGSSENFYSVFEALSNNEIPVIKASLAGTRIVGRLSVGNKNGL 70
T. thermophila 1 --MARRCQFENSNDIGVFCKLTSAYCLVSVGASENFYSVFESELVPHIPVVIHTSGGTRIVGRVTCGNKNGL 70
S. cerevisiae 1 --MATRTQFENSNEIGVFSKLTNTYCLVAVGGSSENFYSVFEALGDAPIPVHTTLAGTRIGRMTAGNRRGL 70
D. melanogaster 1 --MALRVQFENNDDIGVFTKLTNTYCLVAIGGSETFYSAFAEALGDTIPVVHANVGGCRIGRLTVGNRNL 70
C. elegans 1 --MALRVDYEGSNDVGVFCTLNSYCLVGVGGTQNFYSILEAELSDLIPVVHTSASTRIVGRLLTVGNRHGL 70
H. sapiens 1 --MAVRASFENNCEIGCFACKLTNTYCLVAIGGSENFYSVFEGLSDTIPVHVASAGCRIGRMVCGNRHGL 70
```

```
M. thermautotrophicus 69 VVSNQAMDRIDALAAAG-----VEAVRIPERFTAVGNLVLVANDNGAVASPLSDDALQVIGDVLVVD-VK 133
M. jannaschii 69 LLPKIVEDEELDRIKNFLKKNLNDLVEIISKNTALGNLILITNDKALISPEL-KDFKKDIEDSLNVE-VE 140
A. fulgidus 66 VVSDQVLSSLR---ELEKS----LDVLIETPMTFCFNNLLINDRGGIANPEMSSVVEKVAFMDEI-LV 129
P. horikoshii 70 ILPWYWDAAIERIKKALSEYGDMEVVPFRSKYTALGNLILITNDKAAALVSAKFSRKEAQEICEI LGVE-VE 140
S. acidocaldarius 70 LLPRTALADEIKVIEKQAKD---VRVEVVDIRPTALGNLILITNDKALIQDLSAEAINKVKKALQIDTAI 137
D. discoideum 71 LLPNTCTDQELQQRNSLDP--DVVVQRIIEEKSALGNLCIATNDYVALVHPDIDRETEEIIADVLGVE-VF 138
Z. mays 71 LLPHTTTDQELQHLRNSLDP--QVVVQRIDERLSALGNLCIACNDHVALTHPDLDKETEEFIADVLGVE-VF 138
O. lucimarinus 71 LLPTSATDQVMHIRNSLDP--EVIQRTEERLSALGNLCVACNDHVALVHPDIDQETEEIIADVLGVE-VF 138
T. thermophila 71 LVPNTCNDNLRNIRNSLDP--NVRVRRIEEKL SALGNLCVANDYVALIHPDLDRSEEIIADVLGVE-VF 138
S. cerevisiae 71 LVPNTCTDQELQHLRNSLDP--SVKIQRVEERLSALGNVICNDYVALVHPDIDRETEEIIADVLGVE-VF 138
D. melanogaster 71 LVPNSTTDQELQHLRNSLDP--AVKIYRVEERLSALGNVICNDYVALVHPDLDKETEEIIADVLKVE-VF 138
C. elegans 71 LVPNATTDQELQHLRNSLDP--EVAIRVDERLSALGNVICNDHVALVHAEISAEETEALVEVLKVE-VF 138
H. sapiens 71 LVPNNTTDQELQHIRNSLDP--TVQIRRVEERLSALGNVITTCNDYVALVHPDLDRTEEIIADVLKVE-VF 138
```

```
M. thermautotrophicus 134 VSTLAGLNIVGSMGAATNRGALLNPQASSEIIGITIEDTLGVLEAD-VGTVNHGVTLLIGACSVANSNGVLVGE 204
M. jannaschii 141 IGTIAELPTVGSNAVVTNKGCLTHPLVEDDELEFLKSLFKVEYIGKCTANKGTTSSVGACIIANSKGAUVVGGD 212
A. fulgidus 130 KGTVGGIKTVMGAAVVTNRGGLANPNINNEWAKKLEQVAGVEVL-TGTVFGTDMVGSGLVANSKGVVVGGRD 200
P. horikoshii 141 RGVIAGLHAVGSAGVVTNKGGLVHPPTSEDELEWLSDLFKVDVY-VGTANMCPVYVGTCLMANRNGVVGHL 211
S. acidocaldarius 138 KGTIANIITVGSVAVITDKAGLVHIDATEEELKLSLSELFKVKLD-SGTVFGSVFIRSGLVANSRNGVLVGS 208
D. discoideum 139 RQTVSGNVLVGTICALTNQAGLVHPMPTSIADQDELSLQLQVPLV-AGTVNRGNECVAAAGCVVNDWTAIVGAD 209
Z. mays 139 RQTIAGNILVGSYCAFNSNRDGLVHPHTSVEDLDELSTLQVPLV-AGTVNRGSEVIAAGMTVNDWTAFCGLD 209
O. lucimarinus 139 RQSVAGNVLVGSFCKFSNRGGMVHPQASVQEVDELSSLQVPLV-AGTVNRGSDVIGAGLLVNDWIAFCGLD 209
T. thermophila 139 RQTISGNILVGSYCSLSNQGGLVHPPTSVQDQEESSLLQVPLV-AGTVNRGSDVIGAGLVNDWIAFCGLD 209
S. cerevisiae 139 RQTISGNILVGSYCSLSNQGGLVHPPTSVQDQEESSLLQVPLV-AGTVNRGSDVIGAGLVNDWIAFCGLD 209
D. melanogaster 139 RQTIAQNSLVGSYVLSNQGGMVHPKTSIQDQDELSLQVPLV-AGTVNRGSEVLAAGMVDWLSFVGMN 209
C. elegans 139 RVTLAQNSLVGSYVLSNQGGLVAARTPPETQREIAALLQIPVV-AGTCNRGSELIGAGMVDWVAFVCGLD 209
H. sapiens 139 RQTVADQVVLVGSYVLSNQGGLVHPKTSIQDQDELSLQVPLV-AGTVNRGSEVIAAGMVDWCAFCGLD 209
```

```
M. thermautotrophicus 205 TIGPELARIIEALGFLEG----- 222
M. jannaschii 213 TIGPELLIIEDALGLI----- 228
A. fulgidus 201 TIGFELGIVEEALFP----- 215
P. horikoshii 212 TIGPEIVKIEALGFV----- 227
S. acidocaldarius 209 TIGAEILRIQRAFSD----- 223
D. discoideum 210 TITATEISVIESIFALQG-SKPSNIIINNIRNSIVDNV- 244
Z. mays 210 TITATELSVIESVFKLRE-GQPTAIVEDMRKSLIDSYV 245
O. lucimarinus 210 TITATEIQVIESIYQLHG-ADSSGAVRDIRASLLDTLV 245
T. thermophila 210 TITATEISVVENIFKLNEMKDENMDNEMRKDFVMNLE 245
S. cerevisiae 210 TITATELSVIESIFRLQD-AQPESISGNLRDTLLETYS 245
D. melanogaster 210 TITATEISVIESVFKLNQ-AQPATVTTKLRALLIEDMS 245
C. elegans 210 TITATELSVIESIFKLGEGCAPTSISNQLRDTLLETSM 246
H. sapiens 210 TITATELSVIESVFKLNE-AQPSTIATSMRDSLIDSLT 245
```

(b)

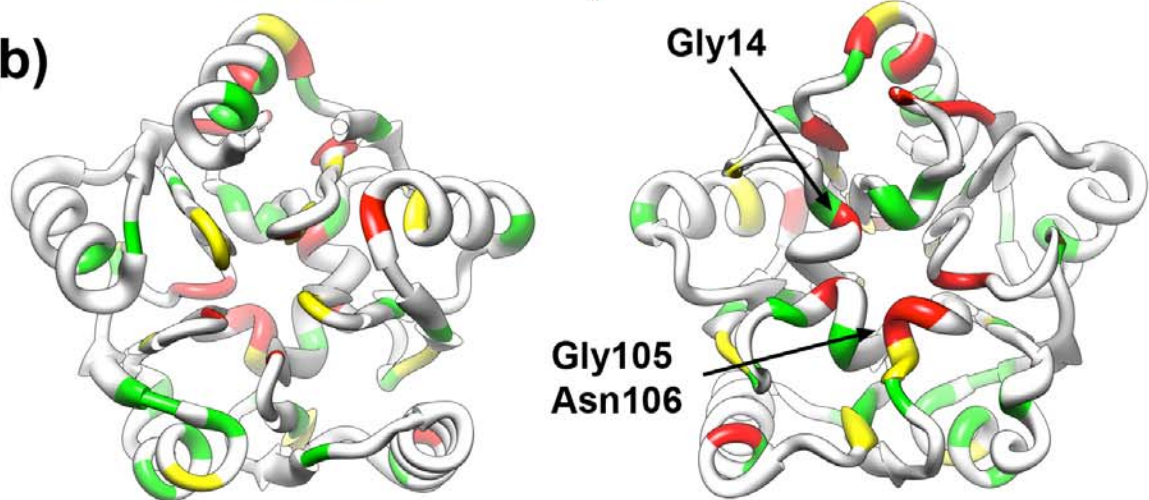
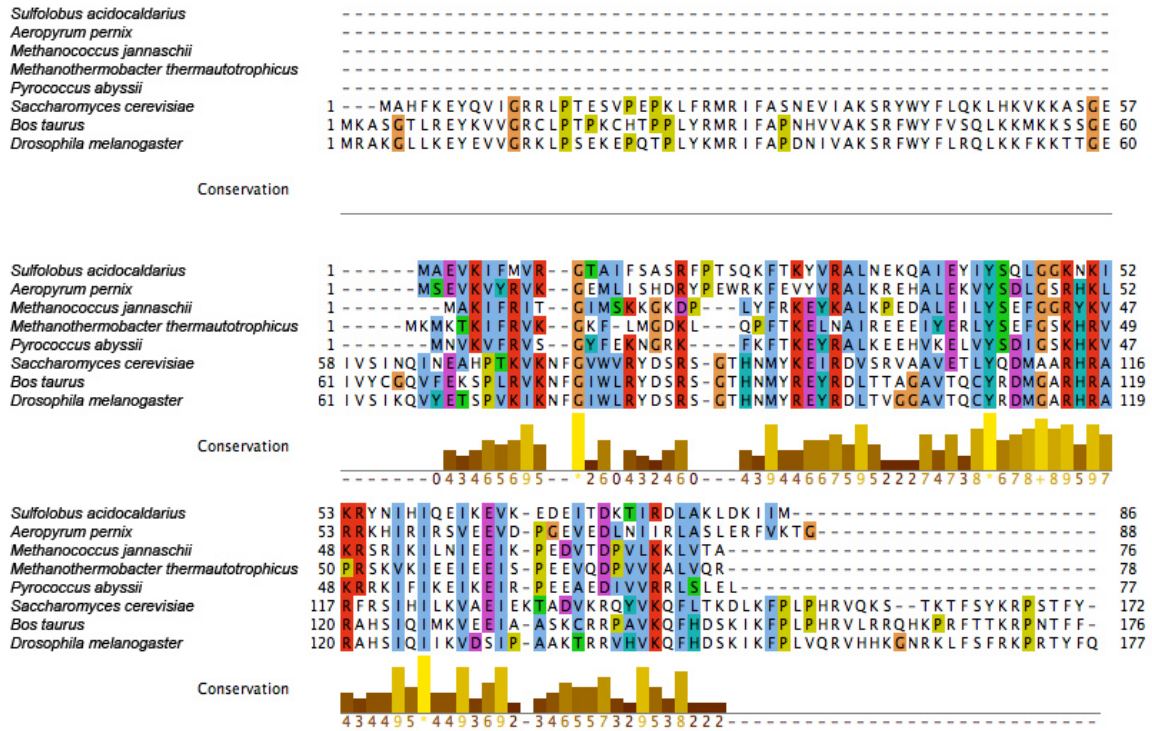


Fig. S3. (a) Multiple sequence alignment of IF6 protein sequences from selected eukaryotic and archaeal organisms. The incompletely conserved phosphorylatable Serine residues 174, 175, and 235^{15;16;17} are marked with red asterisks. Phosphorylation of Ser235 is controversial

^{17; 18}. (b) Ribbon diagrams of the *M. jannaschii* aIF6 X-ray crystal structure (PDB ID: **1G61**¹⁹) (left: solvent side; right: ribosome interaction surface) colored according to Clustal characters (red: identity according to (a); green, yellow, white: conservation in descending order). The highly conserved Gly14, Gly105, and Asn106 (*S. cerevisiae* nomenclature) are discussed in the supplementary text (see above).

Fig. S4. Aligned 23S rRNA sequences of *M. thermautotrophicus* (*Mt*) and *H. marismortui* (*Hm*) extracted from a multiple sequence alignment of archaeal 23S rRNA²⁰ (Comparative RNA Website, <http://www.rna.cccb.utexas.edu>). 23S rRNA expansions in *M. thermautotrophicus* are indicated, and labelled with the designation of the corresponding eukaryotic rRNA sequence element. A red asterisk next to the segment designation indicates that the corresponding expansion is of sufficient size to be interpreted by docking of a model into the *M. thermautotrophicus* 50S cryo-EM map.

(a)



(b)

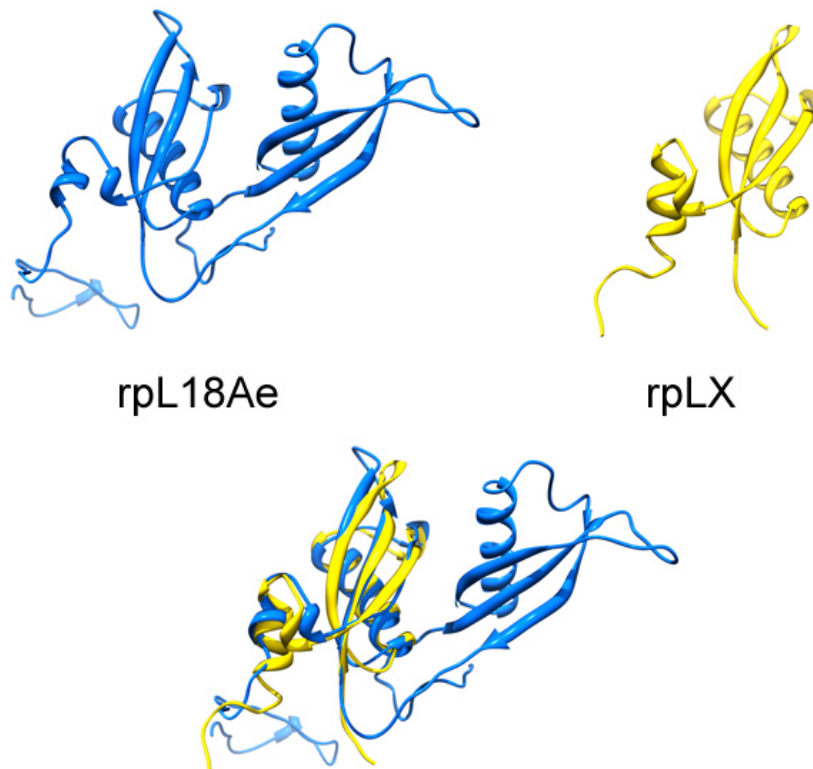


Fig. S5. (a) Multiple sequence alignment showing the homology between rpL18Ae (eukaryotes: *S. cerevisiae*, *B. taurus*, *D. melanogaster*) and rpLX (archaea: *S. acidocaldarius*,

A. permix, *M. jannaschii*, *M. thermautotrophicus*, *P. abyssii*). (b) Comparison of the X-ray crystal structure of ribosome-bound rpL18Ae²¹ (*T. thermophila*; blue) and the solution structure of rpLX (*M. thermautotrophicus*, PDB ID: **2JXT**; yellow).

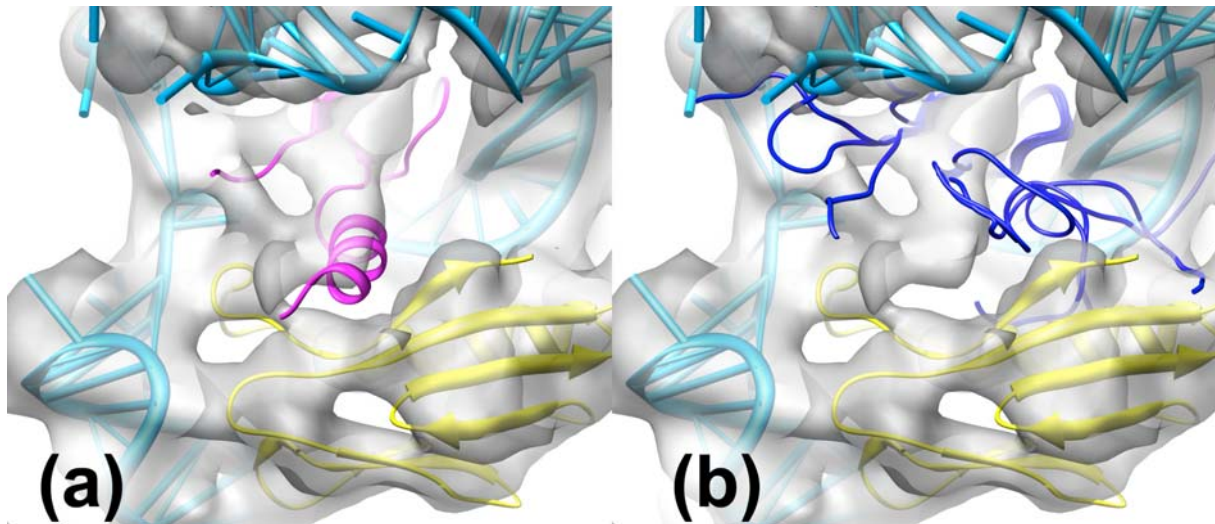


Fig. S6. Ribosomal protein rpL40e may fold fully only after making contact with ribosomal RNA. (a) *M. thermautotrophicus* rpL40e (pink) modelled according to the ribosome-bound *T. thermophila* template²¹ shown in the 50S-aIF6 density. (b) Solution structure of *S. solfataricus* rpL40e (dark blue; PDB ID: **2AYJ**, chains A-D) aligned onto the *Mt* rpL40e core fold. The density features are consistent with formation of an α -helix at the rpL40e N-terminus upon binding to the ribosomal subunit. Ribosomal protein rpL6p in pale yellow, 23S rRNA in sky blue.

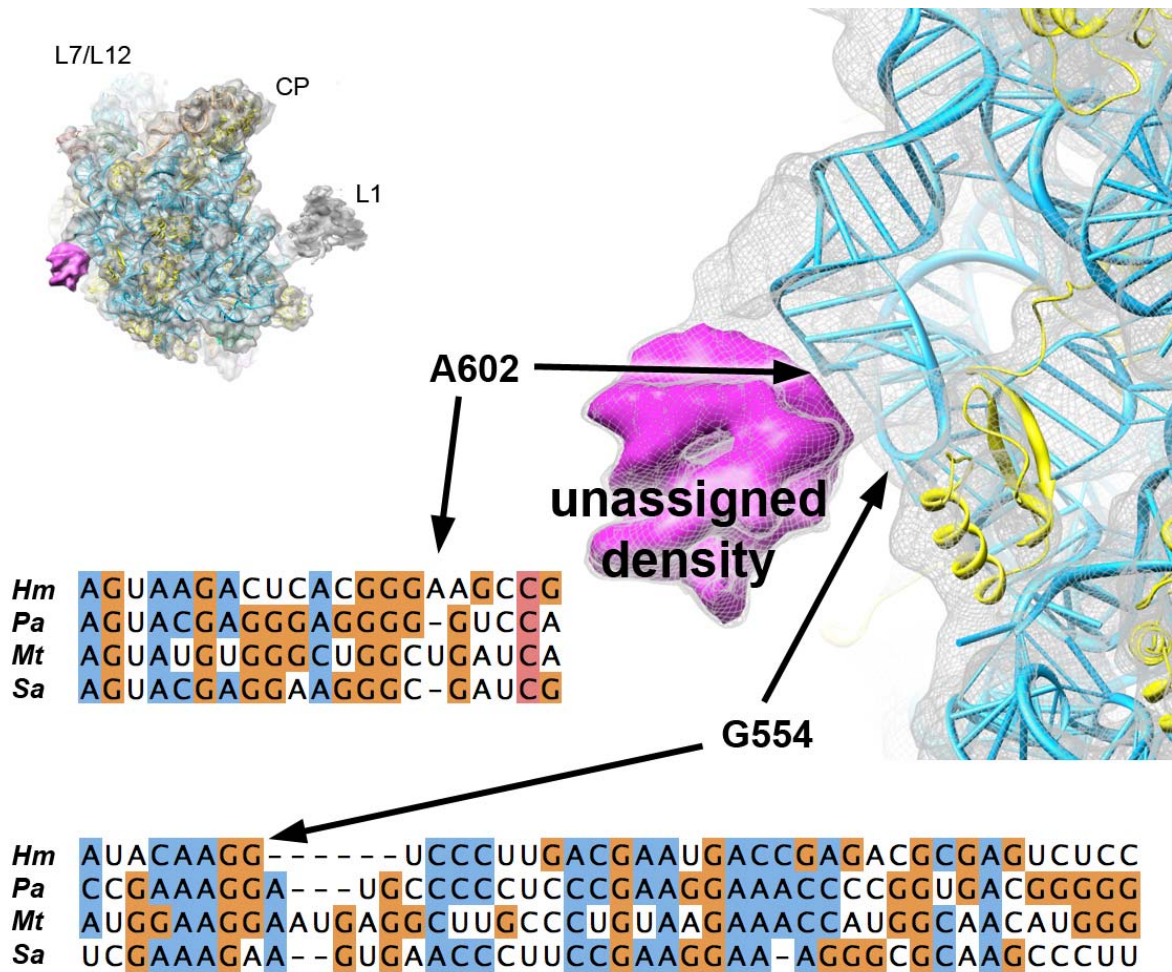


Fig. S7. The unassigned density on the solvent side of the *M. thermautotrophicus* 50S subunit corresponds most likely to an unidentified protein. No rRNA expansions of sufficient size are present in the immediate vicinity, and the variability of rRNA in this region does not correlate with the presence of this density in *S. acidocaldarius* (*Sa*) and *M. thermautotrophicus* (*Mt*) and its absence in *P. aerophilum* (*Pa*) and *H. marismortui* (*Hm*). The inset shows the 50S-aIF6 complex in the same orientation as the main figure. The L1 and L7/L12 stalks and the central protuberance (CP) are labeled. The sequence alignment was taken from a multiple sequence alignment of archaeal 23S rRNA²⁰ (Comparative RNA Website, <http://www.rna.ccbb.utexas.edu>). A new alignment of only the four displayed sequences using ClustalX²² resulted in only minor changes of the alignment (not shown) and did not affect the conclusion. The depiction of the molecular model and the alignment were created in Chimera²³ and JalView 2.4²⁴, respectively.

Table S1. Overview of the *M. thermotrophicus* (*Mt*) large ribosomal subunit protein analysis by LC-MS/MS (green: protein found; red: protein not found).

<i>Mt</i> ribosomal protein	Molecular weight (kDa)	LC-MS/MS	<i>Mt</i> ribosomal protein	Molecular weight (kDa)	LC-MS/MS
rpL3P	37.3	Green	rpL18e	13.5	Green
rpLP0/rpL10P	36.5	Green	rpL7Ae	13.2	Green
rpL4P	28.3	Green	rpL32e	12.6	Green
rpL2P	26.1	Green	rpL21e	11.1	Green
rpL1P	23.8	Green	rpL44e	10.9	Green
rpL15e	21.9	Green	rpL12P	10.5	Green
rpL18P	21.6	Green	rpL30e	10.5	Green
rpL6P	19.8	Green	rpL34e	10.3	Red
rpL5P	19.4	Green	rpL23P	9.9	Green
rpL10e	18.1	Green	rpL37Ae	9.9	Red
rpL11p	17.5	Green	rpL31e	9.4	Green
rpL22P	17.3	Green	rpLX	9.3	Green
rpL30P	17.1	Green	rpL14e	8.3	Green
rpL19e	17.1	Green	rpL29P	7.4	Green
rpL15P	16.5	Green	rpL37e	7.1	Red
rpL13P	16.1 ^a	Green	rpL39e	6.4	Green
rpL14P	14.2	Green	rpL24e	6.3	Red
rpL24P	13.9	Green	rpL40e	5.6	Green

^a estimate based on alignment of rpL13P/rpS9 fusion protein annotated in *M. thermotrophicus* genome with *M. marburgensis* homolog

Table S2. Assignment of the rotational orientation of aIF6 on the 50S ribosomal subunit by cross-correlation. The cross-correlation was computed using the fit in map command in Chimera, using a map simulated from the *M. thermautotrophicus* aIF6 homology model at 6.7 Å resolution for docking into the cryo-EM density. The rotational position is measured clockwise relative to the position in the 60S-eIF6 complex ²¹.

Position	Rotation	Correlation
1	0°	0.82
2	72°	0.78
3	144°	0.77
4	216°	0.80
5	288°	0.76

Table S3. Overview of the *H. marismortui* and *M. thermautotrophicus* large ribosomal subunit protein content (SwissProt database²⁵). Proteins that were not present in the *H. marismortui* crystal structure²⁶ but present in the *H. marismortui* genome are indicated in orange, while proteins absent from the *H. marismortui* genome are indicated in green.

<i>M. thermautotrophicus</i> ribosomal protein	<i>H. marismortui</i> ribosomal protein
rpL4P	rpL4P
rpL22P	rpL22P
rpL3P	rpL3P
rpL24P	rpL24P
rpL2P	rpL2P
rpL14P	rpL14P
rpL15P	rpL15P
rpL18P	rpL18P
rpL1P	rpL1P
rpL23P	rpL23P
rpL24e	rpL24e
rpL30P	rpL30P
rpL5P	rpL5P
rpL6P	rpL6P
rpL12P	rpL12P
rpLX	rpLX
rpL13P	rpL13P
rpL10e	rpL10e
rpL11p	rpL11p
rpL14e	-
rpL15e	rpL15e
rpL18e	rpL18e
rpL19e	rpL19e
rpL21e	rpL21e
rpL29P	rpL29P
rpL30e	-

rpL31e	rpL31e
rpL32e	rpL32e
rpL34e	-
rpL37Ae	rpL37Ae
rpL37e	rpL37e
rpL39e	rpL39e
rpL44e	rpL44e
rpL7Ae	rpL7Ae
rpL40e	rpL40e
rpLP0	rpL10P / rpL10E / rplP0

Table S4. Generation of the *M. thermautotrophicus* (*Mt*) 50S-aIF6 protein model based on *H. marismortui* (*Hm*), *T. thermophila* (*Tt*), and *M. jannaschii* (*Mj*) protein structures^{19; 21; 26}.

<i>Mt</i> ribosomal protein	Template ribosomal protein	Source structure for <i>Mt</i> model	Residues used in <i>Mt</i> model	Full length <i>Mt</i> protein	Chain ID in <i>Mt</i> model
rpL2p	Hm rpL2p	3CC2.A	1-238	1-241	A
rpL3p	Hm rpL3p	3CC2.B	1-338	1-337	B
rpL4p	<i>Hm</i> rpL4p	3CC2.C	1-246	1-254	C
rpL5p	<i>Hm</i> rpL5p	3CC2.D	11-176	1-168	D
rpL6p	<i>Hm</i> rpL6p	3CC2.E	1-173	1-177	E
rpL7Ae	<i>Hm</i> rpL7Ae	3CC2.F	1-120	1-123	F
rpL10e	<i>Hm</i> rpL10e	3CC2.H	8-176	1-160	H
rpL13p	<i>Hm</i> rpL13p	3CC2.J	7-145	1-145 ^a	J
rpL14p	<i>Hm</i> rpL14p	3CC2.K	1-132	1-132	K
rpL15p	<i>Hm</i> rpL15p	3CC2.L	1-145	1-146	L
rpL15e	<i>Hm</i> rpL15e	3CC2.M	4-185	1-182	M
rpL18p	<i>Hm</i> rpL18p	3CC2.N	1-185	1-192	N
rpL18e	<i>Hm</i> rpL18e	3CC2.O	1-116	1-121	O
rpL19e	<i>Hm</i> rpL19e	3CC2.P	2-149	1-148	P
rpL21e	<i>Hm</i> rpL21e	3CC2.Q	1-96	1-96	Q
rpL22p	<i>Hm</i> rpL22p	3CC2.R	1-151	1-153	R
rpL23p	<i>Hm</i> rpL23p	3CC2.S	1-85	1-86	S
rpL24p	<i>Hm</i> rpL24p	3CC2.T	1-117	1-117	T
rpL24e	<i>Hm</i> rpL24e	3CC2.U	4-55	1-53	U
rpL29p	<i>Hm</i> rpL29p	3CC2.V	1-64	1-64	V
rpL30p	<i>Hm</i> rpL30p	3CC2.W	1-154	1-152	W
rpL31e	<i>Hm</i> rpL31e	3CC2.X	7-92	1-81	X
rpL32e	<i>Hm</i> rpL32e	3CC2.Y	125-232	1-108	Y
rpL37Ae	<i>Hm</i> rpL37Ae	3CC2.Z	25-114	1-89	Z
rpL37e	<i>Hm</i> rpL37e	3CC2.1	1-57	1-60	1
rpL39e	<i>Hm</i> rpL39e	3CC2.2	1-50	1-51	2

rpL44e	<i>Hm</i> rpL44e	3CC2.3	1-92	1-92	3
rpLX	<i>Mt</i> rpLX	2JXT.A	1-78	1-78	G
rpL40e	<i>Tt</i> rpL40e	4A19.K	80-129	1-48	5
rpL14e	<i>Tt</i> rpL14e	4A19.F	4-80	1-75	7
rpL30e	<i>Tt</i> rpL30e	4A19.G	8-103	1-98	6
rpL34e	<i>Tt</i> rpL34e	4A19.L	8-97	1-88	4
aIF6	<i>Mj</i> aIF6 ^b	1G61.A	3-227 (+1)	1-222	I

^a estimate based on alignment of rpL13P/rpS9 fusion protein annotated in

M. thermotrophicus genome with *M. marburgensis* homolog

^b model contains *M. thermotrophicus* protein sequence

Table S5. Generation of the *M. thermautotrophicus* (*Mt*) 50S-aIF6 rRNA model based on *H. marismortui* (*Hm*), *T. thermophila* (*Tt*), and *S. cerevisiae* (*Sc*) rRNA structures^{21; 26; 27}.

<i>Mt</i> rRNA	Template rRNA	Source structure for <i>Mt</i> model	Residues used in <i>Mt</i> model	Chain ID in <i>Mt</i> model
5S rRNA	Hm 5S rRNA	3CC2.9	1-122	9
23S rRNA	Hm 23S rRNA	3CC2.0	10-2914 ^a	0
ES3	<i>Tt</i> 5.8S rRNA	4A17.2	115-134	8
ES5	<i>Tt</i> 26S rRNA	4A19.1	125-144	8
ES20	<i>Tt</i> 26S rRNA	4A19.1	1645-1678	8
ES24	<i>Sc</i> 25S rRNA	3O58.1	1753-1772	8
ES26	<i>Tt</i> 26S rRNA	4A19.1	1827-1849	8
ES41	<i>Tt</i> 26S rRNA	4A19.1	3302-3317	8

^a Contains gaps to accommodate the insertions of expansion segments ES3, ES5, ES20, ES24, ES26, and ES41.

Supplementary References

1. Menne, T. F., Goyenechea, B., Sánchez-Puig, N., Wong, C. C., Tonkin, L. M., Ancliff, P. J., Brost, R. L., Costanzo, M., Boone, C. & Warren, A. J. (2007). The Shwachman-Bodian-Diamond syndrome protein mediates translational activation of ribosomes in yeast. *Nat Genet* **39**, 486-95.
2. Lecompte, O., Ripp, R., Thierry, J.-C., Moras, D. & Poch, O. (2002). Comparative analysis of ribosomal proteins in complete genomes: an example of reductive evolution at the domain scale. *Nucleic Acids Research* **30**, 5382-90.
3. Cole, S. T., Eiglmeier, K., Parkhill, J., James, K. D., Thomson, N. R., Wheeler, P. R., Honoré, N., Garnier, T., Churcher, C., Harris, D., Mungall, K., Basham, D., Brown, D., Chillingworth, T., Connor, R., Davies, R. M., Devlin, K., Duthoy, S., Feltwell, T., Fraser, A., Hamlin, N., Holroyd, S., Hornsby, T., Jagels, K., Lacroix, C., Maclean, J., Moule, S., Murphy, L., Oliver, K., Quail, M. A., Rajandream, M. A., Rutherford, K. M., Rutter, S., Seeger, K., Simon, S., Simmonds, M., Skelton, J., Squares, R., Squares, S., Stevens, K., Taylor, K., Whitehead, S., Woodward, J. R. & Barrell, B. G. (2001). Massive gene decay in the leprosy bacillus. *Nature* **409**, 1007-11.
4. Katinka, M. D., Duprat, S., Cornillot, E., Méténier, G., Thomarat, F., Prensier, G., Barbe, V., Peyretailade, E., Brottier, P., Wincker, P., Delbac, F., El Alaoui, H., Peyret, P., Saurin, W., Gouy, M., Weissenbach, J. & Vivarès, C. P. (2001). Genome sequence and gene compaction of the eukaryote parasite *Encephalitozoon cuniculi*. *Nature* **414**, 450-3.
5. Hershberg, R., Tang, H. & Petrov, D. A. (2007). Reduced selection leads to accelerated gene loss in *Shigella*. *Genome Biol* **8**, R164.
6. Andersson, S. G., Zomorodipour, A., Andersson, J. O., Sicheritz-Pontén, T., Alsmark, U. C., Podowski, R. M., Näslund, A. K., Eriksson, A. S., Winkler, H. H. & Kurland, C. G. (1998). The genome sequence of *Rickettsia prowazekii* and the origin of mitochondria. *Nature* **396**, 133-40.
7. Moran, N. A., McLaughlin, H. J. & Sorek, R. (2009). The dynamics and time scale of ongoing genomic erosion in symbiotic bacteria. *Science* **323**, 379-82.
8. Waters, E., Hohn, M. J., Ahel, I., Graham, D. E., Adams, M. D., Barnstead, M., Beeson, K. Y., Bibbs, L., Bolanos, R., Keller, M., Kretz, K., Lin, X., Mathur, E., Ni, J., Podar, M., Richardson, T., Sutton, G. G., Simon, M., Soll, D., Stetter, K. O., Short, J. M. & Noordewier, M. (2003). The genome of *Nanoarchaeum equitans*: insights into early archaeal evolution and derived parasitism. *Proc Natl Acad Sci USA* **100**, 12984-8.
9. Rocap, G., Larimer, F. W., Lamerdin, J., Malfatti, S., Chain, P., Ahlgren, N. A., Arellano, A., Coleman, M., Hauser, L., Hess, W. R., Johnson, Z. I., Land, M., Lindell, D., Post, A. F., Regala, W., Shah, M., Shaw, S. L., Steglich, C., Sullivan, M. B., Ting, C. S., Tolonen, A., Webb, E. A., Zinser, E. R. & Chisholm, S. W. (2003). Genome divergence in two *Prochlorococcus* ecotypes reflects oceanic niche differentiation. *Nature* **424**, 1042-7.
10. Dufresne, A., Garczarek, L. & Partensky, F. (2005). Accelerated evolution associated with genome reduction in a free-living prokaryote. *Genome Biol* **6**, R14.
11. Smole, Z., Nikolic, N., Supek, F., Šmuc, T., Sbalzarini, I. F. & Krisko, A. (2011). Proteome sequence features carry signatures of the environmental niche of prokaryotes. *BMC Evol Biol* **11**, 26.
12. Kennedy, S. P., Ng, W. V., Salzberg, S. L., Hood, L. & DasSarma, S. (2001). Understanding the adaptation of *Halobacterium* species NRC-1 to its extreme environment through computational analysis of its genome sequence. *Genome Research* **11**, 1641-50.

13. Gandbhir, M., Rasched, I., Marlière, P. & Mutzel, R. (1995). Convergent evolution of amino acid usage in archaeobacterial and eubacterial lineages adapted to high salt. *Res Microbiol* **146**, 113-20.
14. Rosenthal, P. B. & Henderson, R. (2003). Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. In *Journal of Molecular Biology*, Vol. 333, pp. 721-45.
15. Basu, U., Si, K., Deng, H. & Maitra, U. (2003). Phosphorylation of mammalian eukaryotic translation initiation factor 6 and its *Saccharomyces cerevisiae* homologue Tif6p: evidence that phosphorylation of Tif6p regulates its nucleocytoplasmic distribution and is required for yeast cell growth. *Mol Cell Biol* **23**, 6187-99.
16. Biswas, A., Mukherjee, S., Das, S., Shields, D., Chow, C. W. & Maitra, U. (2011). Opposing action of casein kinase 1 and calcineurin in nucleo-cytoplasmic shuttling of mammalian translation initiation factor eIF6. *J Biol Chem* **286**, 3129-38.
17. Ceci, M., Gaviraghi, C., Gorrini, C., Sala, L. A., Offenhäuser, N., Marchisio, P. C. & Biffo, S. (2003). Release of eIF6 (p27BBP) from the 60S subunit allows 80S ribosome assembly. *Nature* **426**, 579-84.
18. Finch, A. J., Hilcenko, C., Basse, N., Drynan, L. F., Goyenechea, B., Menne, T. F., González Fernández, A., Simpson, P., D'Santos, C. S., Arends, M. J., Donadieu, J., Bellanné-Chantelot, C., Costanzo, M., Boone, C., McKenzie, A. N., Freund, S. M. V. & Warren, A. J. (2011). Uncoupling of GTP hydrolysis from eIF6 release on the ribosome causes Shwachman-Diamond syndrome. *Genes & Development* **25**, 917-29.
19. Groft, C. M., Beckmann, R., Sali, A. & Burley, S. K. (2000). Crystal structures of ribosome anti-association factor IF6. *Nat Struct Biol* **7**, 1156-64.
20. Cannone, J. J., Subramanian, S., Schnare, M. N., Collett, J. R., D'Souza, L. M., Du, Y., Feng, B., Lin, N., Madabusi, L. V., Müller, K. M., Pande, N., Shang, Z., Yu, N. & Gutell, R. R. (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.
21. Klinge, S., Voigts-Hoffmann, F., Leibundgut, M. A., Arpagaus, S. & Ban, N. (2011). Crystal Structure of the Eukaryotic 60S Ribosomal Subunit in Complex with Initiation Factor 6. *Science* **334**, 941-8.
22. Thompson, J. D., Gibson, T. J. & Higgins, D. G. (2002). Multiple sequence alignment using ClustalW and ClustalX. *Curr Protoc Bioinformatics* **Chapter 2**, Unit 2.3.
23. Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C. & Ferrin, T. E. (2004). UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* **25**, 1605-12.
24. Waterhouse, A. M., Procter, J. B., Martin, D. M. A., Clamp, M. & Barton, G. J. (2009). Jalview Version 2--a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25**, 1189-91.
25. UniProtConsortium. (2011). Ongoing and future developments at the Universal Protein Resource. *Nucleic Acids Research* **39**, D214-9.
26. Blaha, G., Gürel, G., Schroeder, S. J., Moore, P. B. & Steitz, T. A. (2008). Mutations outside the anisomycin-binding site can make ribosomes drug-resistant. In *Journal of Molecular Biology*, Vol. 379, pp. 505-19.
27. Ben-Shem, A., Jenner, L., Yusupova, G. & Yusupov, M. (2010). Crystal structure of the eukaryotic ribosome. *Science* **330**, 1203-9.