

Supplementary Materials

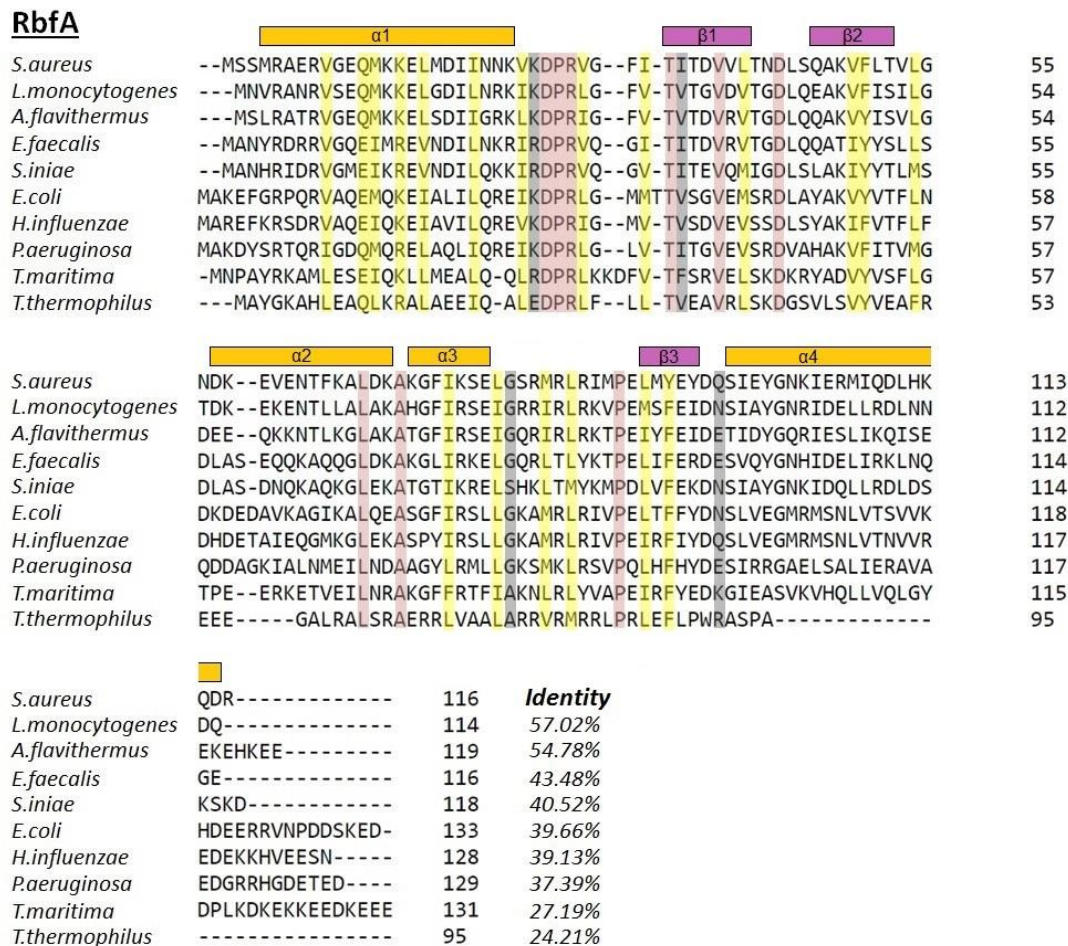
Yet another similarity between mitochondrial and bacterial ribosomal small subunit biogenesis obtained by structural characterization of RbfA from *S. aureus*

Aydar G. Bikmullin, Bulat Fatkhullin, Artem Stetsenko, Azat Gabdulkhakov, Natalia Garaeva, Liliia Nurullina, Evelina Klochkova, Alexander Golubev, Iskander Khusainov, Natalie Trachtmann, Dmitriy Blokhin, Albert Guskov, Shamil Validov, Konstantin Usachev and Marat Yusupov

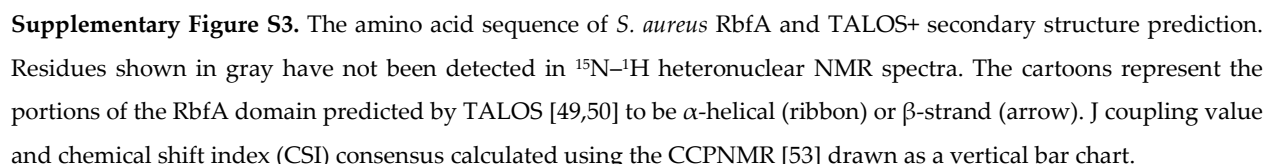
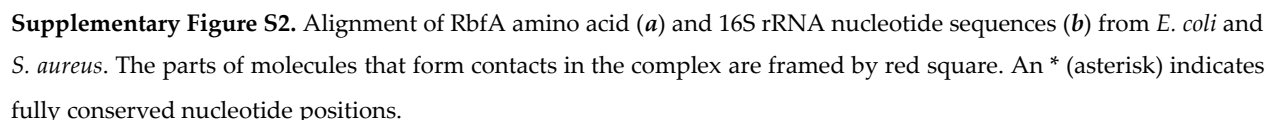
This file includes:

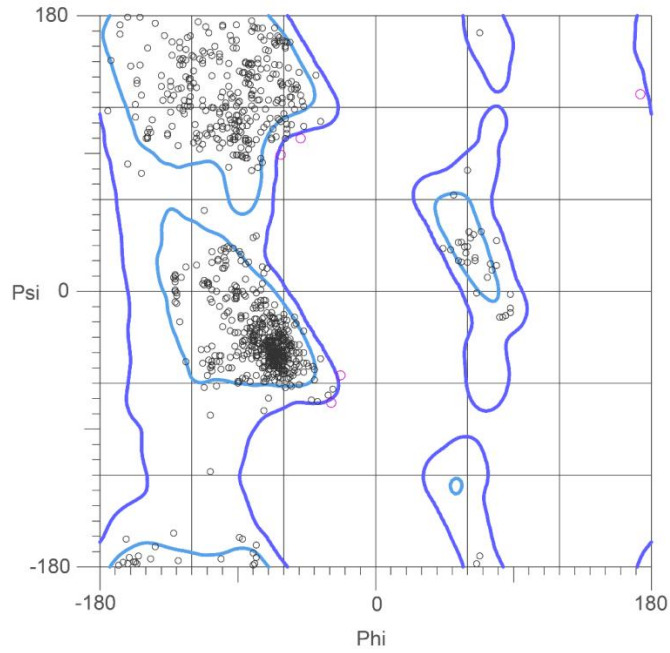
- I. Supplementary Figures S1, S2, S3, S4, S5, S6, S7, S8, S9 with legend;
- II. Supplementary Tables S1, S2, S3, S4.

I. SUPPLEMENTARY FIGURES



Supplementary Figure S1. Multiple sequence alignment and identity percentages of RbfA from 10 different bacterial species performed in Clustal Omega [58]. Pink shading indicates positions that have a single, fully conserved residue (identity), yellow – conservation between amino acids with substantial sequence identity, and gray – amino acids with low sequence identity. The secondary structure elements of *S. aureus* RbfA are framed in orange and violet.

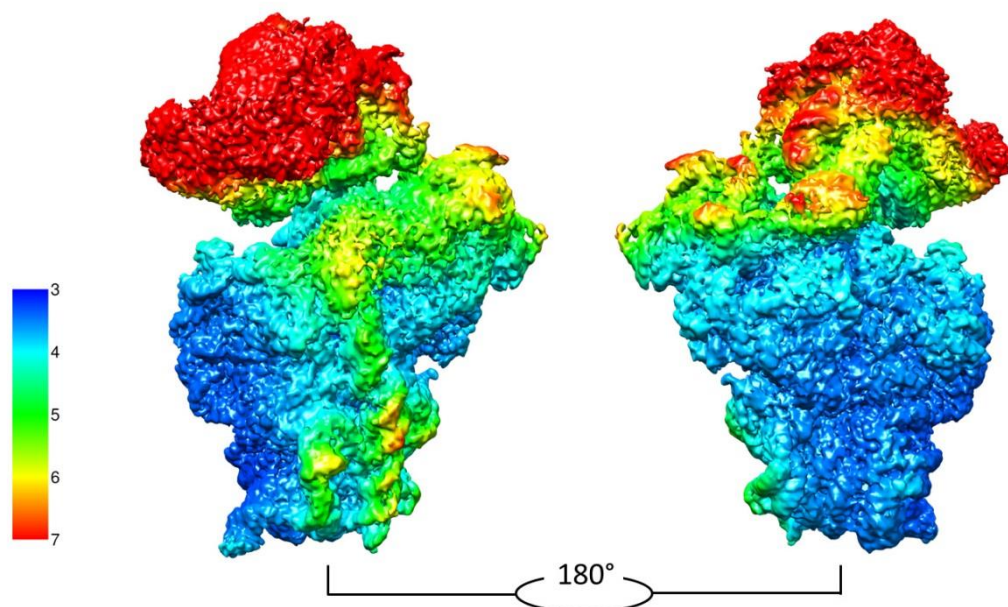




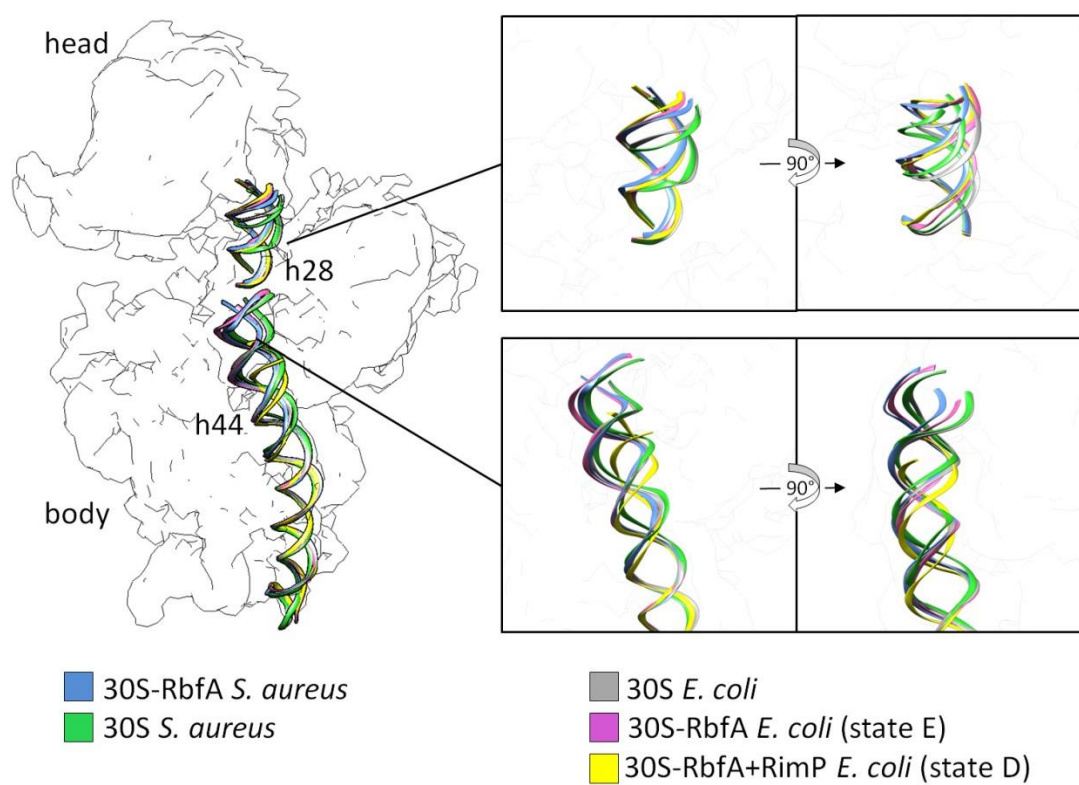
Supplementary Figure S4. The Ramachandran plot of *S. aureus* RbfA conformers.

<p>#1</p> <p>Seq: GSSMRAERVG EQMKKELMDI INNKKVDPRV GFITITDVVL TNDLSQAKVF 50</p> <p>SS: cccHHHHHH HHHHHHHHH HHHHScTTS ccEEEEEE cTTSccEEEE 100</p> <p>Seq: LTVLGNDKEV ENTFKALDKA KGFIKSELGS RMRLRIMPEL MYEYDQSIIEY 100</p> <p>SS: EEESScHHHH HHHHHHHHH HHHHHHHHHT TcScSSccEE EEEccSSSc</p> <p>Seq: GNKIERMIQD LHKQDRVE 118</p> <p>SS: cccScccc cGGGccc</p>	<p>#6</p> <p>Seq: GSSMRAERVG EQMKKELMDI INNKKVDPRV GFITITDVVL TNDLSQAKVF 50</p> <p>SS: cccHHHHHH HHHHHHHHH HHHHScTTS ccEEEEEE cTTSccEEEE 100</p> <p>Seq: LTVLGNDKEV ENTFKALDKA KGFIKSELGS RMRLRIMPEL MYEYDQSIIEY 100</p> <p>SS: EEESScHHHH HHHHHHHHH HHHHHHHHHT SSccSSccEE EEEccSSSc</p> <p>Seq: GNKIERMIQD LHKQDRVE 118</p> <p>SS: cccccScSc cTTGGGc</p>
<p>#2</p> <p>Seq: GSSMRAERVG EQMKKELMDI INNKKVDPRV GFITITDVVL TNDLSQAKVF 50</p> <p>SS: cccHHHHHH HHHHHHHHH HHHHScTTS ccEEEEEE cTTSccEEEE 100</p> <p>Seq: LTVLGNDKEV ENTFKALDKA KGFIKSELGS RMRLRIMPEL MYEYDQSIIEY 100</p> <p>SS: EEcSScHHHH HHHHHHHHH HHHHHHHHHT TcScSccEE EEEccSSSc</p> <p>Seq: GNKIERMIQD LHKQDRVE 118</p> <p>SS: cccScccc cGGGccc</p>	<p>#7</p> <p>Seq: GSSMRAERVG EQMKKELMDI INNKKVDPRV GFITITDVVL TNDLSQAKVF 50</p> <p>SS: cccHHHHHH HHHHHHHHH HHHHScTTS ccEEEEEE cTTSccEEEE 100</p> <p>Seq: LTVLGNDKEV ENTFKALDKA KGFIKSELGS RMRLRIMPEL MYEYDQSIIEY 100</p> <p>SS: EEESScHHHH HHHHHHHHH HHHHHHHHHT cccSSccEE EEEccSSSc</p> <p>Seq: GNKIERMIQD LHKQDRVE 118</p> <p>SS: cccSccSc cTTcSc</p>
<p>#3</p> <p>Seq: GSSMRAERVG EQMKKELMDI INNKKVDPRV GFITITDVVL TNDLSQAKVF 50</p> <p>SS: cccHHHHHH HHHHHHHHH HHHHScTTS ccEEEEEE cTTSccEEEE 100</p> <p>Seq: LTVLGNDKEV ENTFKALDKA KGFIKSELGS RMRLRIMPEL MYEYDQSIIEY 100</p> <p>SS: EEESScHHHH HHHHHHHHH HHHHHHHHHT ScScSccEE EEEccSSSc</p> <p>Seq: GNKIERMIQD LHKQDRVE 118</p> <p>SS: ccTccBecc BGGGccc</p>	<p>#8</p> <p>Seq: GSSMRAERVG EQMKKELMDI INNKKVDPRV GFITITDVVL TNDLSQAKVF 50</p> <p>SS: cccHHHHHH HHHHHHHHH HHHHScTTS ScEEEEEE cSScEEEE 100</p> <p>Seq: LTVLGNDKEV ENTFKALDKA KGFIKSELGS RMRLRIMPEL MYEYDQSIIEY 100</p> <p>SS: EEcScHHHH HHHHHHHHH HHHHHHHHHT SSccSSccEE EEEccTTT 118</p> <p>Seq: GNKIERMIQD LHKQDRVE 118</p> <p>SS: cccSScccc cSGGccc</p>
<p>#4</p> <p>Seq: GSSMRAERVG EQMKKELMDI INNKKVDPRV GFITITDVVL TNDLSQAKVF 50</p> <p>SS: cccHHHHHH HHHHHHHHH HHHHScTTS ccEEEEEE cTTSccEEEE 100</p> <p>Seq: LTVLGNDKEV ENTFKALDKA KGFIKSELGS RMRLRIMPEL MYEYDQSIIEY 100</p> <p>SS: EEESScHHHH HHHHHHHHH HHHHHHHHHT TTcSScEE EEEccSSSc</p> <p>Seq: GNKIERMIQD LHKQDRVE 118</p> <p>SS: cccSccSc cTTTTcc</p>	<p>#9</p> <p>Seq: GSSMRAERVG EQMKKELMDI INNKKVDPRV GFITITDVVL TNDLSQAKVF 50</p> <p>SS: cccHHHHHH HHHHHHHHH HHHHScTTS ScEEEEEE cTTSccEEEE 100</p> <p>Seq: LTVLGNDKEV ENTFKALDKA KGFIKSELGS RMRLRIMPEL MYEYDQSIIEY 100</p> <p>SS: EEESScHHHH HHHHHHHHH HHHHHHHHHT SSccSSccEE EEEccSSSc</p> <p>Seq: GNKIERMIQD LHKQDRVE 118</p> <p>SS: ccSccScSc cTTTTc</p>
<p>#5</p> <p>Seq: GSSMRAERVG EQMKKELMDI INNKKVDPRV GFITITDVVL TNDLSQAKVF 50</p> <p>SS: cccHHHHHH HHHHHHHHH HHHHScTTS ScEEEEEE cTTSccEEEE 100</p> <p>Seq: LTVLGNDKEV ENTFKALDKA KGFIKSELGS RMRLRIMPEL MYEYDQSIIEY 100</p> <p>SS: EEESScHHHH HHHHHHHHH HHHHHHHHHT cccSSccEE EEEccSSSc</p> <p>Seq: GNKIERMIQD LHKQDRVE 118</p> <p>SS: cccccccc cTTTSc</p>	<p>#10</p> <p>Seq: GSSMRAERVG EQMKKELMDI INNKKVDPRV GFITITDVVL TNDLSQAKVF 50</p> <p>SS: cccHHHHHH HHHHHHHHH HHHHScTTS ScEEEEEE cTTSccEEEE 100</p> <p>Seq: LTVLGNDKEV ENTFKALDKA KGFIKSELGS RMRLRIMPEL MYEYDQSIIEY 100</p> <p>SS: EEcScHHHH HHHHHHHHH HHHHHHHHHT SSccSSccEE EEEccSSSc</p> <p>Seq: GNKIERMIQD LHKQDRVE 118</p> <p>SS: ccSScccc cSSSc</p>

Supplementary Figure S5. The *S. aureus* RbfA secondary structure assignment by the DSSP program [34] for region 95D-118E (yellow) in several NMR structures. H – α -helix; B – β -bridge; E – β -strand; G – 310 helix; I – π -helix; T – turn; S – bend; C – coil.



Supplementary Figure S6. Local resolution map of *S. aureus* 30S-RbfA complex.



Supplementary Figure S7. Comparison of 16S rRNA h28 and h44 helices of *S. aureus* and *E. coli* in different states (mature or bound with RbfA). Mature *S. aureus* 30S subunit (PDB ID: 5ND8), mature *E. coli* 30S subunit (PDB ID: 7BOD), *E. coli* 30S subunit assembly complex state E (PDB ID: 7BOH), *E. coli* 30S subunit assembly complex state D 79 (PDB ID: 7NAV).

CLUSTAL O(1.2.4) multiple sequence alignment

S.aureus	-----	0
E.coli	-----	0
M.musculus mt	MAAPALRAPLRNSGLALGVRCAVNLPLGTQVRGSRVAFEPREPLIDKEYVRKPAELTE	60
H.sapiens mt	MAAPAVKVARGWNSGLALGVRRAVLQPLGTQVRNSRYSPEFKDPLIDKEYVRKPAELTE	60
T.thermophilus	-----	0
H.influenzae	-----	0
M.pneumoniae	-----	0
T.maritima	-----	0

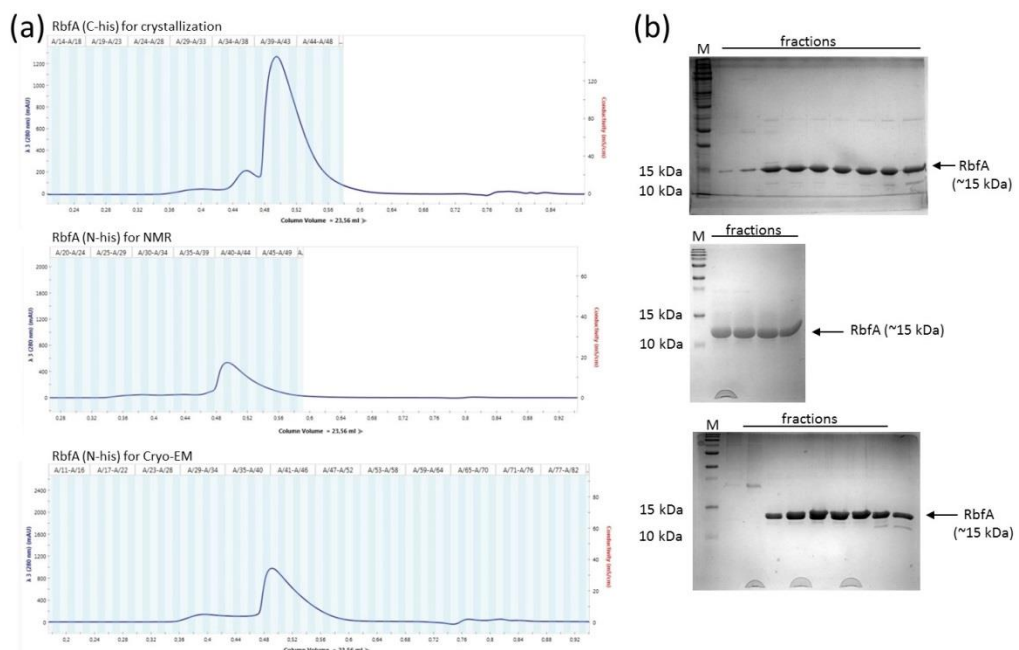
S.aureus	-----HPRKG-SVPKRDVLPDPIHNSKLVTKLINKIMLDGKRGTARILYSAFDLVEQ	52
E.coli	-----HPRRR-VIGQRKILPDPIFGSELLAKFVNILHVDGKSTAESIVYSALETLAQ	52
M.musculus mt	EEKYDQELKKQFIKAAATSTSVFADPVISKFTNMMHGGKVLARSLSMAQTLEAVR	120
H.sapiens mt	EEKYVRELKKQFIKAAAGKTSVFEDPVISKFTNMMHGGKVLARSLSMAQTLEAVR	120
T.thermophilus	-----HARRR-RAEVRLQPDLYGDLVLTAFINKIMRDGKKLAARIFYDACKITQE	52
H.influenzae	-----HPRRR-SVEARKILPDPIFGSELLAKFVNILHVDGKSTAESIVYSALETLAQ	52
M.pneumoniae	-----HRRN-RAPKRTVLDPVFNITLVTRIINVIHMDGKKGLAQRLVGFADLIEQ	51
T.maritima	-----HRRR-RAEKRIIPDPVFGDLVAKLINRVHMDGKKTAQRIYVGFADLIRE	51

S.aureus	-----RSGRDALVFEEAINIMPVLEKARRVGGSNVQVPEVRPERRT	97
E.coli	-----RSKGSELEAFEALENVRPTVEKSRVGGSTYQVPEVRPVRNR	97
M.musculus mt	KQFEKYRAASAEQATIERNPYRIFHEALKNCEPVLGLVPLTKGGHFYQVPLADRRRR	180
H.sapiens mt	KQFEKYRAASAEQATIERNPYRIFHEALKNCEPVLGLVPLTKGGHFYQVPLADRRRR	180
T.thermophilus	-----KTGQELKVFQAVENKPRMEVRSRRVGGANVQVPEVSPRRQQ	97
H.influenzae	-----RTGKELEAFEALENVRPTVEKSRVGGSTYQVPEVRPVRNR	97
M.pneumoniae	-----RTKEKPLTVFERAVGNMPRLERVRRIAGSNVQVPEVPQDRKI	96
T.maritima	-----KTKKDPLEVFRQAVENKPVLEVRPRVGGATYQVPEVQEPRT	96

S.aureus	TLGLRWLVHYARLRGE--KTMEDRLANEILDAANITGGAVKKREDTHMAEANKAFAHYR	155
E.coli	ALAMRWIVEAARKRGD--KSMALRLANESDAENKGTAVKKREDVHMAEANKAFAHYR	155
M.musculus mt	FLAMKWMITECRENKPRTLMPKLSHELLFAFHNRPVIRKHNHMAEANKRALAHYR	240
H.sapiens mt	FLAMKWMITECRENKPRTLMPKLSHELLFAFHNRPVIRKHNHMAEANKRALAHYR	240
T.thermophilus	SLALRWLVQAANRPE--RRAAVRIAEHLMDAEGKGGAVKKEDVERMAEANKRAYAHYR	155
H.influenzae	ALGMRWIVEAARKRGD--KSMALRLANESDAENKGTAVKKREDVHMAEANKAFAHYR	155
M.pneumoniae	ALALRWIAFARKRHE--KTMLEKIANEIIASNNITGAIIKKDDTHMAEANKAFAHYR	154
T.maritima	SLALRWIVEAARAKKG--RPMKEKLAEEIIAAYNNTGTAIIKKEDTHMAEANKAFAHYR	154

S.aureus	-----	156
E.coli	-----LSLRSFHQAGASSKQPALGYLN	179
M.musculus mt	-----	242
H.sapiens mt	-----	242
T.thermophilus	-----	156
H.influenzae	-----	156
M.pneumoniae	-----	155
T.maritima	-----	155

Supplementary Figure S8. Comparison of homologs of S7 ribosomal protein from different organisms was performed in Clustal Omega [58]. Pink shading indicates positions that have a single, fully conserved residue (identity), yellow – conservation between amino acids with substantial sequence identity. The helix of S7 interacting with the C-terminal helix of RbfA is framed in red. An * (asterisk) indicates positions which have a single, fully conserved residue. A : (colon) indicates conservation between groups of strongly similar properties - scoring > 0.5 in the Gonnet PAM 250 matrix. A . (period) indicates conservation between groups of weakly similar properties - scoring <= 0.5 in the Gonnet PAM 250 matrix [58].



II. SUPPLEMENTARY TABLES

Supplementary Table S1. X-ray data collection and refinement statistics.

RbfA	
Resolution range (Å)	32.66 - 2.22 (2.3 - 2.22)
Space group	P 4 ₁
Unit cell	
a, b, c (Å)	65.31, 65.31, 27.28
α , β , γ (°)	90.0, 90.0, 90.0
Multiplicity	7.2 (6.9)
Completeness (%)	99.88 (100.00)
Mean I/sigma(I)	18.60 (2.44)
Wilson B-factor	49.47
R-merge	0.06 (0.78)
R-meas	0.06 (0.85)
R-pim	0.02 (0.32)
CC _{1/2}	0.99 (0.89)
Refinement	
R-work	0.20
R-free	0.24
No. of atoms	852
protein	838
water	14
R.m.s. deviations	
Bond lengths (Å)	0.01
Bond angles (°)	1.01
Ramachandran favored (%)	98.97
Ramachandran allowed (%)	1.03
Clashscore	9.94
Average B-factor	56.79
protein	56.81
water	55.68
Statistics for the highest-resolution shell are shown in parentheses.	

Supplementary Table S2. Statistical information for RbfA structural ensemble.

Distance Constraints	
total	1486
intra-residual	372
inter-residual	1114
sequential (i - j = 1)	495
medium (i - j = 2,3,4)	388
long-range (i - j > 4)	231
Dihedral Angle Constraints	
Phi restraints	85
Psi restraints	85
Structural Statistics	
number of NOE violations	10
number of dihedral angle restraint violations	0
RMSD for bond deviations (Å)	0.01
RMSD for angle deviations (deg)	0.99
RMSD of all backbone atoms (Å)	
Met4-Tyr94	1.02 ± 0.26
Met4-Lys24	0.44 ± 0.12
Asp57-Lys69	0.22 ± 0.11
Lys71-Glu79	0.28 ± 0.09
His112-Asp115	0.75 ± 0.37
Ile33-Leu40	0.34 ± 0.12
Gln46-Val53	0.25 ± 0.09
Glu89-Tyr94	0.30 ± 0.10
Ramachandran Plot	
residues in the most favored region (%)	90.6
residues in the additionally allowed region (%)	8.7
residues in the disallowed region (%)	0.7

Supplementary Table S3. Oligonucleotide primers for *S. aureus* RbfA cloning.

Primer name	Primer sequence
rbfA_Nhis_NdeI_fwd	5'- TTTTTTCATATGAGCAGTATGAGAGCAGAGCGTG T-3'
rbfA_Nhis_HindIII_rev	5'- TTTTTTAAGCTTATCTATCTTGTGTTGTAAATCTT GAATCATTCGTTCAA-3'
rbfA_Chis_NcoI_fwd	5'-TTTTTTCATGAGCAGTATGAGAGCAGAGCG-3'
rbfA_Chis_XhoI_rev	5'- TTTTTCTCGAGTCTATCTTGTGTTGTAAATCTTG AATCATTCGTTCAA-3'

Supplementary Table S4. Data collection, refinement, and validation statistics for *S. aureus* 30S-RbfA complex cryo- EM structure.

<i>Data collection</i>	
Microscope	FEI TALOS ARCTICA
Camera	Gatan K2 Quantum (4k x 4k)
Voltage (kV)	200
Defocus range (μm)	(-1.5) – (-3.0)
Total dose (ē/Å ²)	53.3
Dose per frame(ē/Å ²)	0.89
Number of frames	60
<i>Image processing</i>	
Number of images	876
Number of particles (used for 3D reconstruction)	81421
Resolution (Å; at FSC ^a = 0.143)	2.89
CC ^a (mask)	0.70
CC (box)	0.71
<i>Validation^b</i>	
Clashscore ^c	15.08
<i>Proteins</i>	
MolProbity score	1.97
Poor rotamers (%)	0.93
Favored rotamers (%)	70.51
Ramachandran allowed(%)	4.18
Ramachandran favored (%)	95.82
RMS deviations	
Bonds (Å)	0.005
Angles (°)	0.712
Bad bonds (%)	0.00
Bad angles (%)	0.00
Cis-Prolines (%)	0.00
CaBLAM outliers (%)	3.8
CA Geometry outliers (%)	0.83
<i>RNA</i>	
Probably wrong sugar puckers (%)	0.26
Bad backbone conformations (%)	17.35
Bad bonds (%)	0.00
Bad angles (%)	0.00

^a FSC, Fourier shell correlation; ^b Compiled using MolProbity [46]; CC, correlation coefficient; RMS, root-mean square.