Retapamulin-assisted ribosome profiling reveals the alternative bacterial proteome

Sezen Meydan, James Marks, Dorota Klepacki, Virag Sharma, Pavel Baranov, Andrew Firth, Tõnu Margus, Amira Kefi, Nora Vázquez-Laslop and Alexander S. Mankin

SUPPLEMENTARY INFORMATION

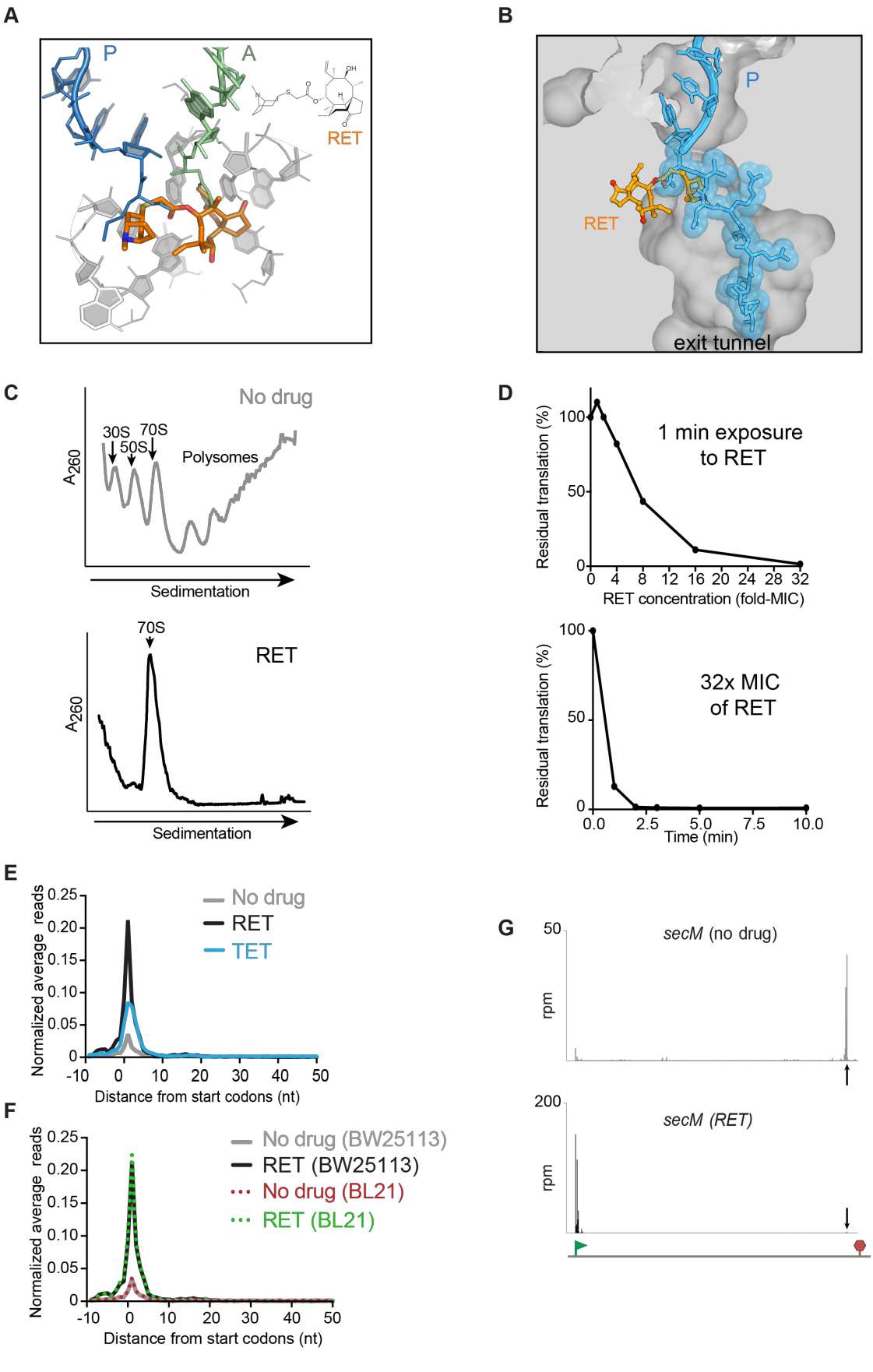

A
P
A
RET
B
P
RET
exit tunnel
C
No drug
30S
70S
50S
Polysomes
A260
Sedimentation
70S
RET
A260
Sedimentation
D
1 min exposure
to RET
Residual translation (%)
100
50
0
0 4 8 12 16 20 24 28 32
RET concentration (fold-MIC)
32x MIC
of RET
Residual translation (%)
100
50
0
0.0 2.5 5.0 7.5 10.0
Time (min)
E
No drug
RET
TET
Normalized average reads
0.25
0.20
0.15
0.10
0.05
0
-10 0 10 20 30 40 50
Distance from start codons (nt)
F
No drug (BW25113)
RET (BW25113)
No drug (BL21)
RET (BL21)
Normalized average reads
0.25
0.20
0.15
0.10
0.05
0
-10 0 10 20 30 40 50
Distance from start codons (nt)
G
50
secM (no drug)
rpm
200
secM (RET)
rpm

Figure S1 Retapamulin arrests ribosomes at initiation, Related to Figure 1

(A) The chemical structure of the pleuromutilin antibiotic retapamulin (RET) bound at the PTC active site of the bacterial ribosome. The model is based on the structural alignment of the 50S ribosomal subunit of *Deinococcus radiodurans* (*Dr*) ribosomes in complex with RET (PDB 2OGO) (Davidovich et al., 2007) and *Thermus thermophilus* 70S ribosomes with fMet-tRNA bound in the P site and Phe-tRNA in the A site (PDB 1VY4) (Polikanov et al., 2014). Note that in the 70S initiation complex, the fMet moiety of the initiator tRNA has to be displaced from the PTC active site to allow for RET binding.

(B) RET cannot coexist with a nascent protein in the ribosome. Alignment of the structures of the *Dr* 50S RET complex with the *E. coli* 70S ribosome carrying ErmBL nascent peptide that esterifies P-site tRNA (PDB 5JTE) (Arenz et al., 2016).

(C) Sucrose gradient analysis of polysome preparation from *E. coli* BW25113 Δ*tolC* cells untreated (top) or treated for 5 min with 12.5 μg/mL (100X MIC) RET. The shown profiles represent cryo-lyzed preparations used in Ribo-seq experiments. Qualitatively similar results have been obtained in analytical experiments with the samples prepared by freezing-thawing (see STAR Methods).

(D) Residual protein synthesis in *E. coli* BL21 Δ*tolC* cells treated with RET, as estimated by incorporation of [^{35}S]-methionine into the TCA-insoluble protein fraction, after 1 min exposure to increasing concentrations of RET (top) or treated with 2 μg/mL of RET (32-fold MIC) for the indicated periods of time (bottom).

(E) Metagene plots comparing the normalized average relative density of ribosomal footprints in *E. coli* BW25113 Δ*tolC* cells untreated (gray trace) or treated 12.5 μg/mL (100X MIC) of RET (black trace). Blue trace represents similar analysis of the publicly-available Ribo-seq data obtained with *E. coli* BW25113 Δ*smpB* cells exposed to tetracycline (TET) [the average of two replicates of Ribo-seq experiments reported in (Nakahigashi et al., 2016)].

(F) Metagene plots comparing the normalized average relative density of ribosomal footprints in the *E. coli* strains BW25113 Δ*tolC* cells or *E. coli* BL21 Δ*tolC* untreated or treated with RET.

(G) Snapshot of ribosomal footprints density in the *secM* gene of *E. coli* BW25113 Δ*tolC* cells untreated or treated with RET. The pTIS and stop codon of the gene are indicated by a green flag and red stop sign, respectively. The black arrow indicates the known site of translation arrest at the codon 165 of the 170-codon *secM* ORF (Nakatogawa and Ito, 2002).

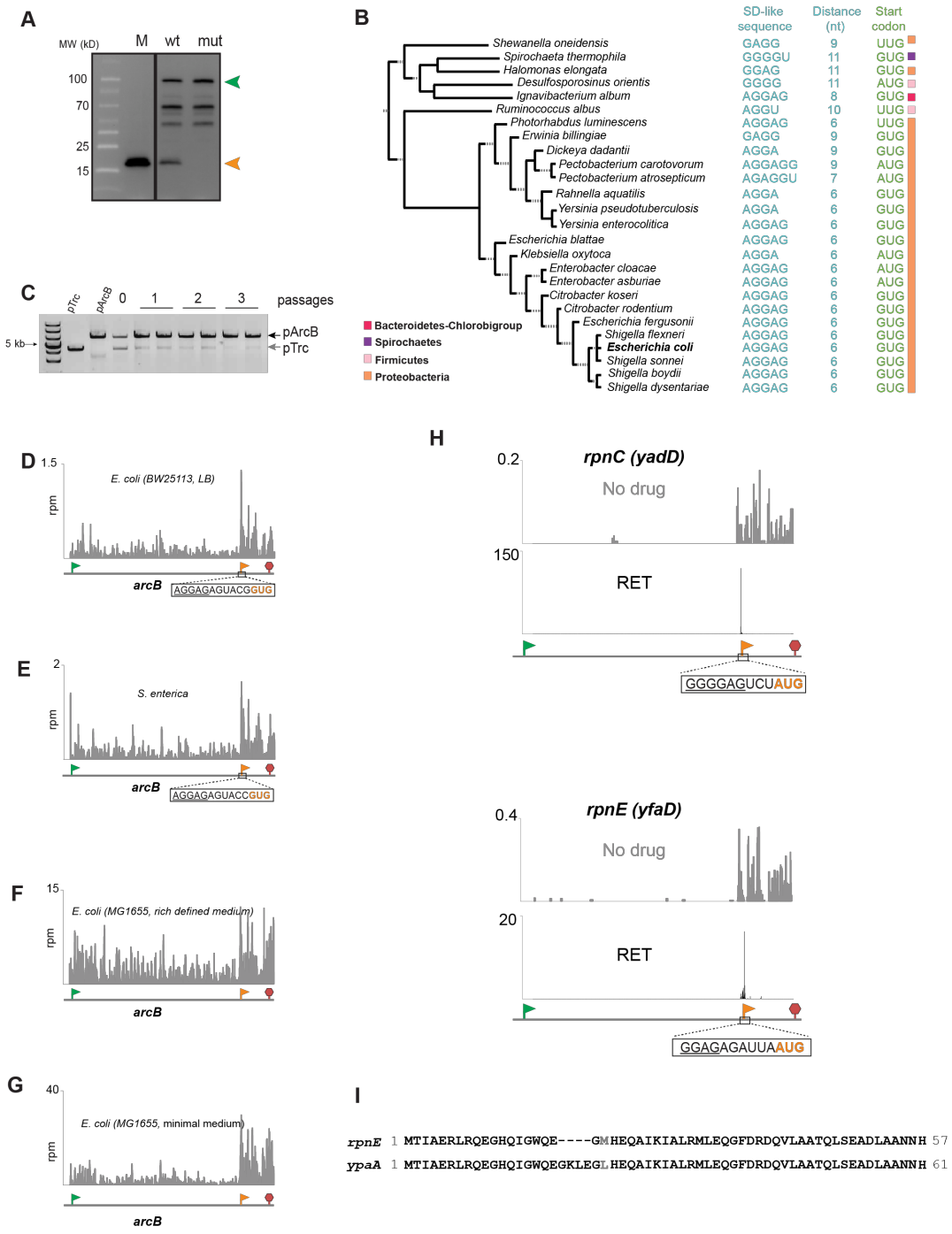


Figure S2 The utilization of an in-frame iTIS within the *arcB* gene leads to production of an alternative protein ArcB-C with a potential role in cell physiology, Related to Figure 3

(A) The uncropped image of the immunoblot shown in Figure 3E, representing the bands corresponding to full-length ArcB-3X FLAG and internal initiation product ArcB-C-3XFLAG (marked with arrow heads). Protein size markers are shown. The origin of the bands marked with dots is unknown.

(B) The iTIS that directs translation of the ArcB-C protein is conserved in the *arcB* gene of diverse bacterial species. The putative start codons and the SD-like sequences are shown.
(C) Presence of *arcB* facilitates *E. coli* growth under low oxygen conditions. BW25113 Δ*arcB E. coli* cells carrying the empty vector pTrc99a or pArcB were co-grown in low oxygen conditions. Gel shows the HindIII-linearized plasmids, isolated from the co-growth cultures to determine fraction of cells with or without *arcB* in the mixture (see Start Methods for details). The ‘’0’ sample represents plasmids from the initial mixture containing equal number of pTrc99A and pArcB cells.
(D-G) The upshift of ribosomal footprints in the *arcB* segment encoding ArcB-C observed in the Ribo-seq profiles of untreated *E. coli* or *Salmonella enterica* cells (Baek et al., 2017; Kannan et al., 2014; Li et al., 2014). The pTIS and iTIS of *arcB* are marked with green and orange flags, respectively, and the stop codon is indicated by a red stop sign.
(G) Representative examples of Ribo-RET and Ribo-seq profiles of two out of five *E. coli rpn* genes.
(H) Alignment of the amino acid sequence of the RpnE-C protein, translated from the iTIS within the *rpnE* gene and the protein encoded in an independent gene *ypaA*.

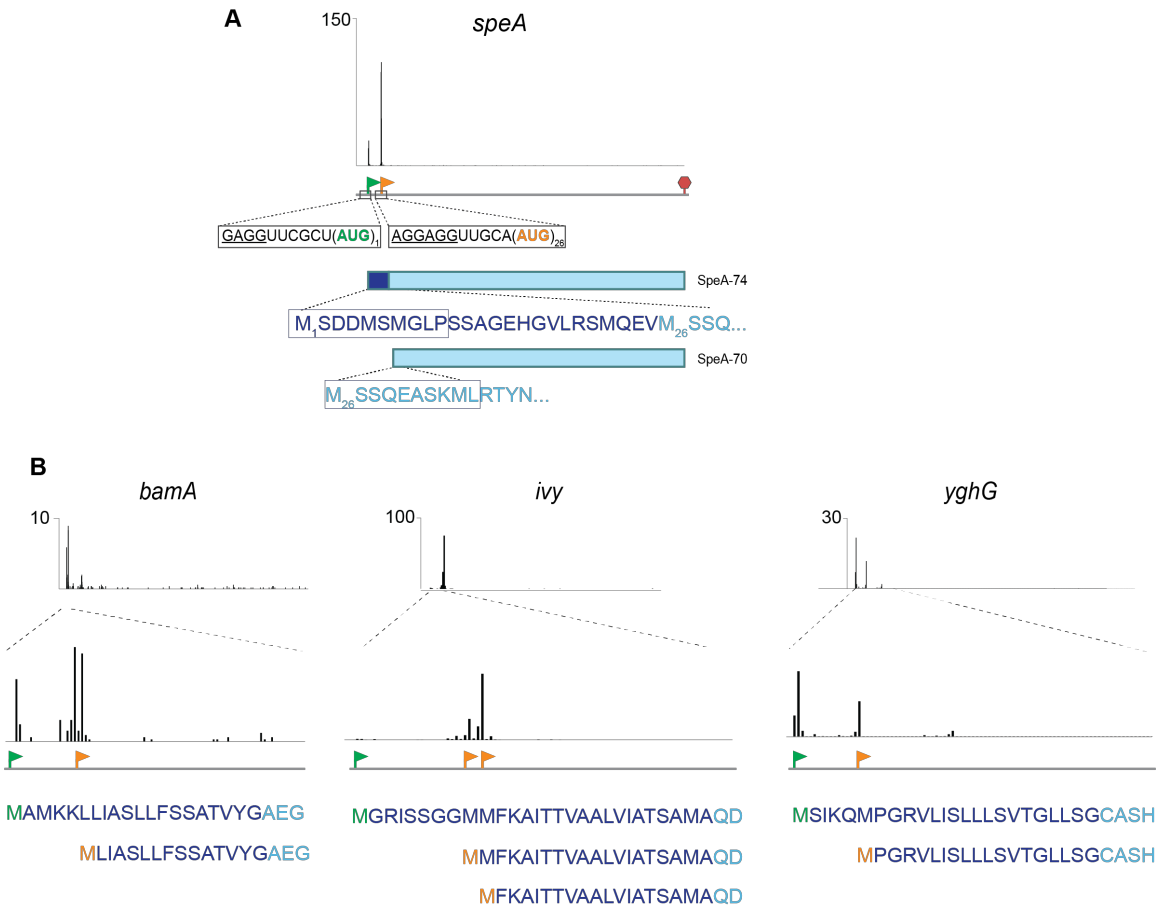


Figure S3. Initiation at the 5'-end proximal iTISs could produce alternative products with incomplete N-terminal signal sequences, Related to Figure 4

(A) Ribo-RET profile of the *speA* gene, showing peaks corresponding to pTIS (green flag) and iTIS (orange flag). The stop codon is indicated by a red stop sign. The putative signal sequence (indicated by dark blue letters) of SpeA-74 (Buch and Boyle, 1985) is lacking in the alternative product SpeA-70 whose translation is initiated at the iTIS. The SpeA isoforms, whose translation is initiated at the pTIS or the iTIS are expected to have different cellular localization. The peptides detected by N-terminomics are boxed (Bienvenut et al., 2015).

(B) Ribo-RET profiles of *bamA, ivy* and *yghG* genes. The N-terminal amino acid sequences of the primary and predicted alternative proteins are indicated. The reported signal sequences are shown in dark blue. The pTISs of the genes are marked by green flags; iTISs are indicated with orange flags.

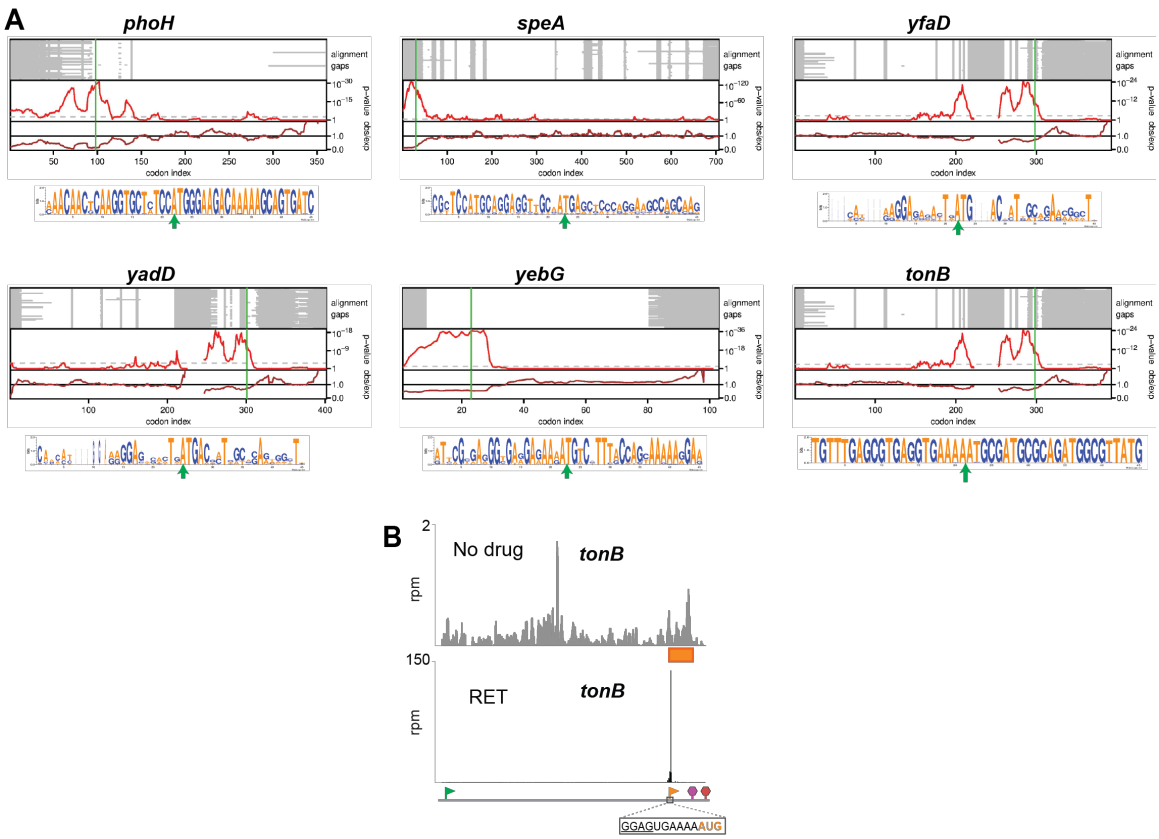


Figure S4. Synonymous site conservation for selected iTISs, Related to Figures 3-5

(A) Synonymous site conservation plots and weblogos for genes with in-frame iTISs (*phoH*, *speA*, *yfaD*, *yadD*, *yebG)* and for the *tonB* gene with an OOF iTIS. Alignment gaps in each sequence are indicated in grey. The two panels show the synonymous substitution rate in a 15-codon sliding window, relative to the CDS average (observed/expected; brown line) and the corresponding statistical significance (*p*-value; red line). The horizontal dashed grey line indicates a *p*-value of 0.05 / (CDS length/window size) – an approximate correction for multiple testing within a single CDS.

(B) An upshift in the local density of ribosome footprints within the alternative frame defined by the *tonB* OOF iTIS (orange rectangle) in cells not exposed to antibiotic. Start codons of the pTIS and OOF iTIS are marked with green and orange flags, respectively, while the respective stop codons are indicated with red and purple stop signs. The start codon and SD-like sequence of the iTIS are shown.

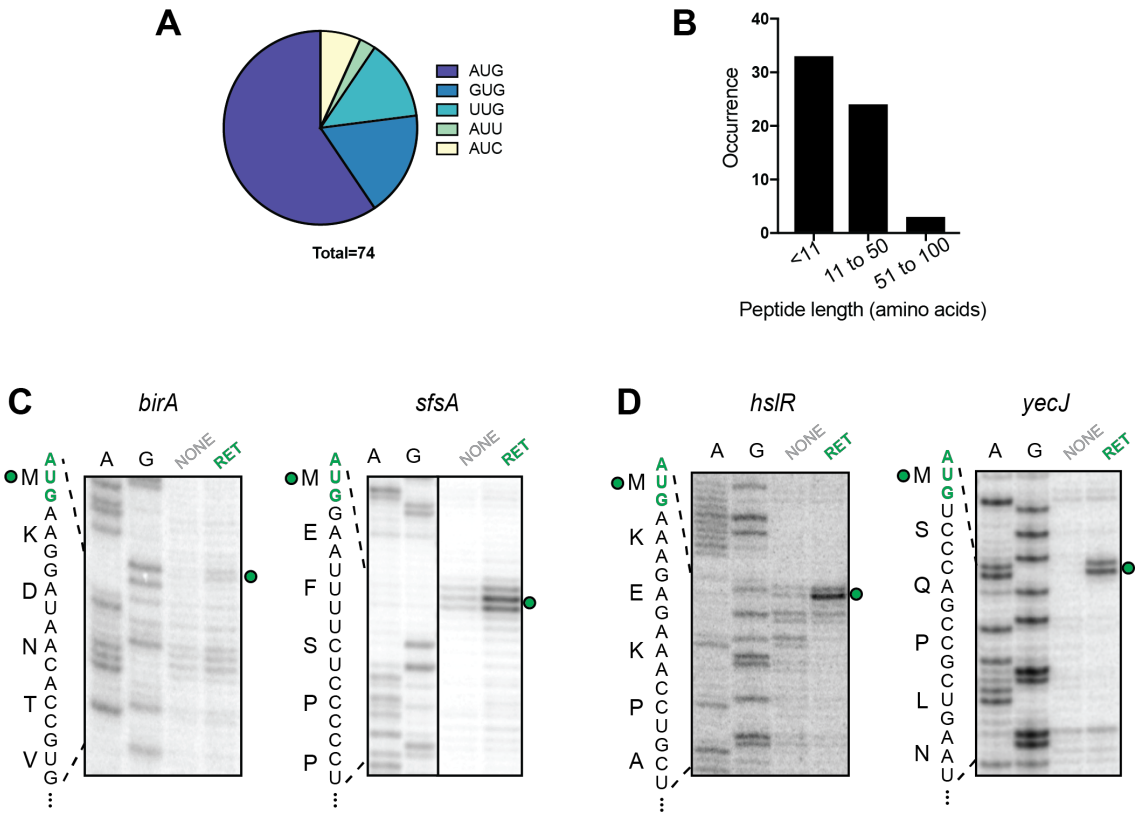


Figure S5 Ribo-RET reveals OOF iTISs, Related to Figures 4 and 5

(A) The distribution of start codons associated with OOF iTISs revealed by Ribo-RET.

(B) The length distribution of the putative alternative proteins whose translation is initiated at OOF iTISs.

(C) and (D) Toe-printing gels showing RET-induced ribosome stalling at the pTISs of *birA* and *sfsA* (shown in Figure 4) and *hslR* and *yecJ* (shown in Figure 5) genes. Samples analyzed in the lanes marked NONE contained no antibiotics. Start codons of the pTISs are indicated in green. Sequencing lanes are shown.

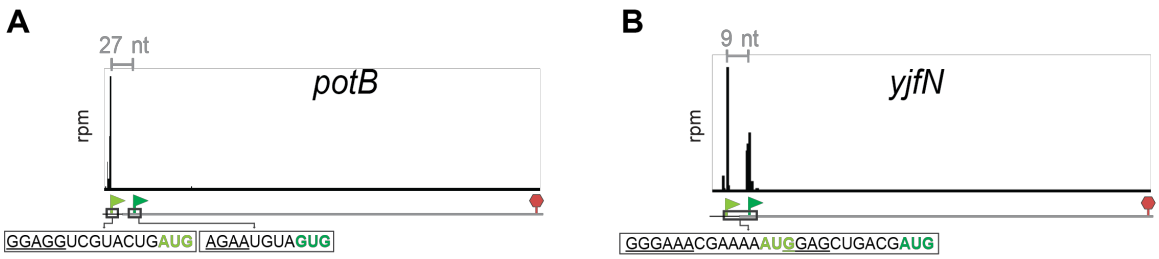
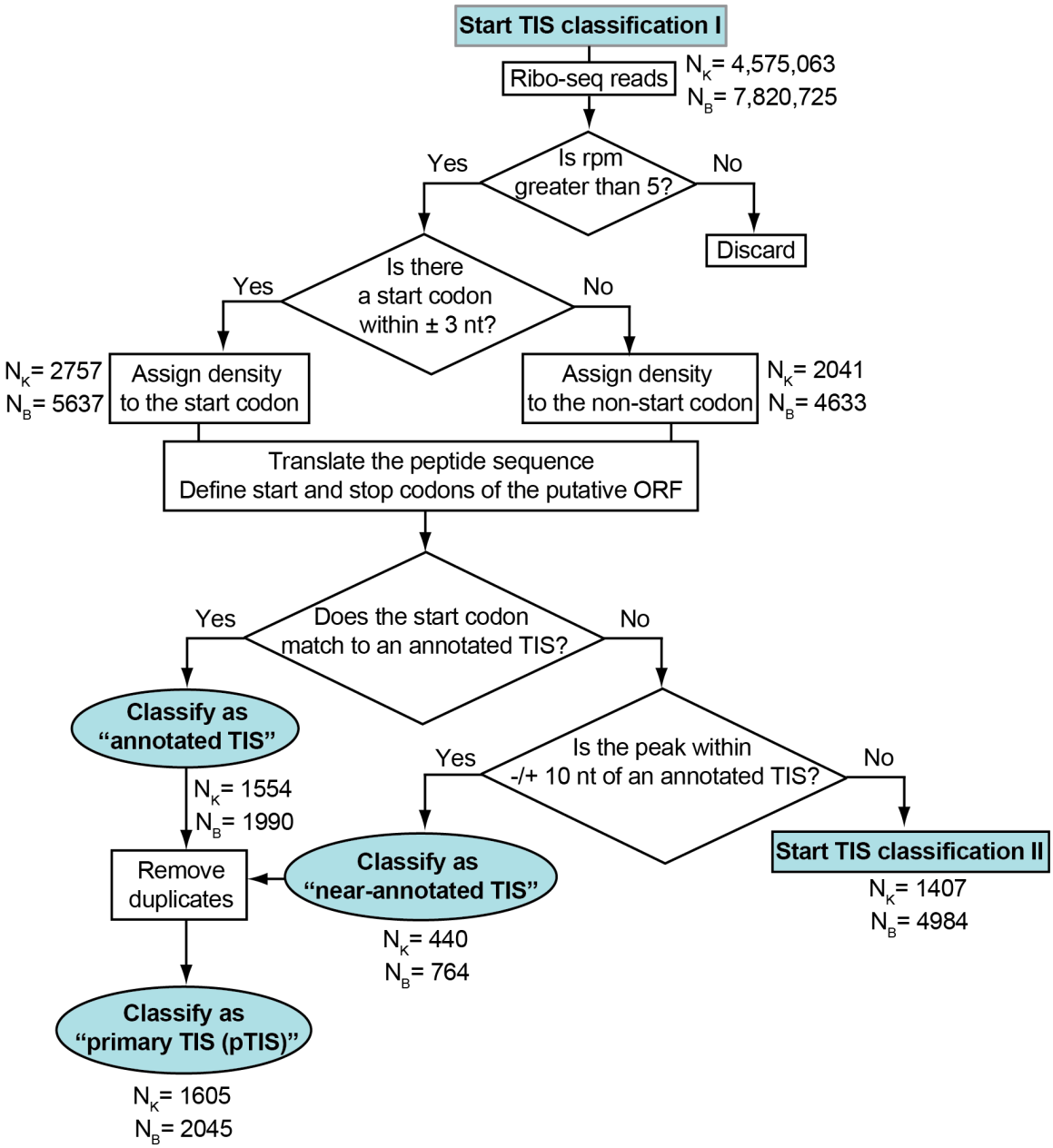


Figure S6. Examples of the genes with Ribo-RET identified TISs outside of the coding regions, Related to Figure 1

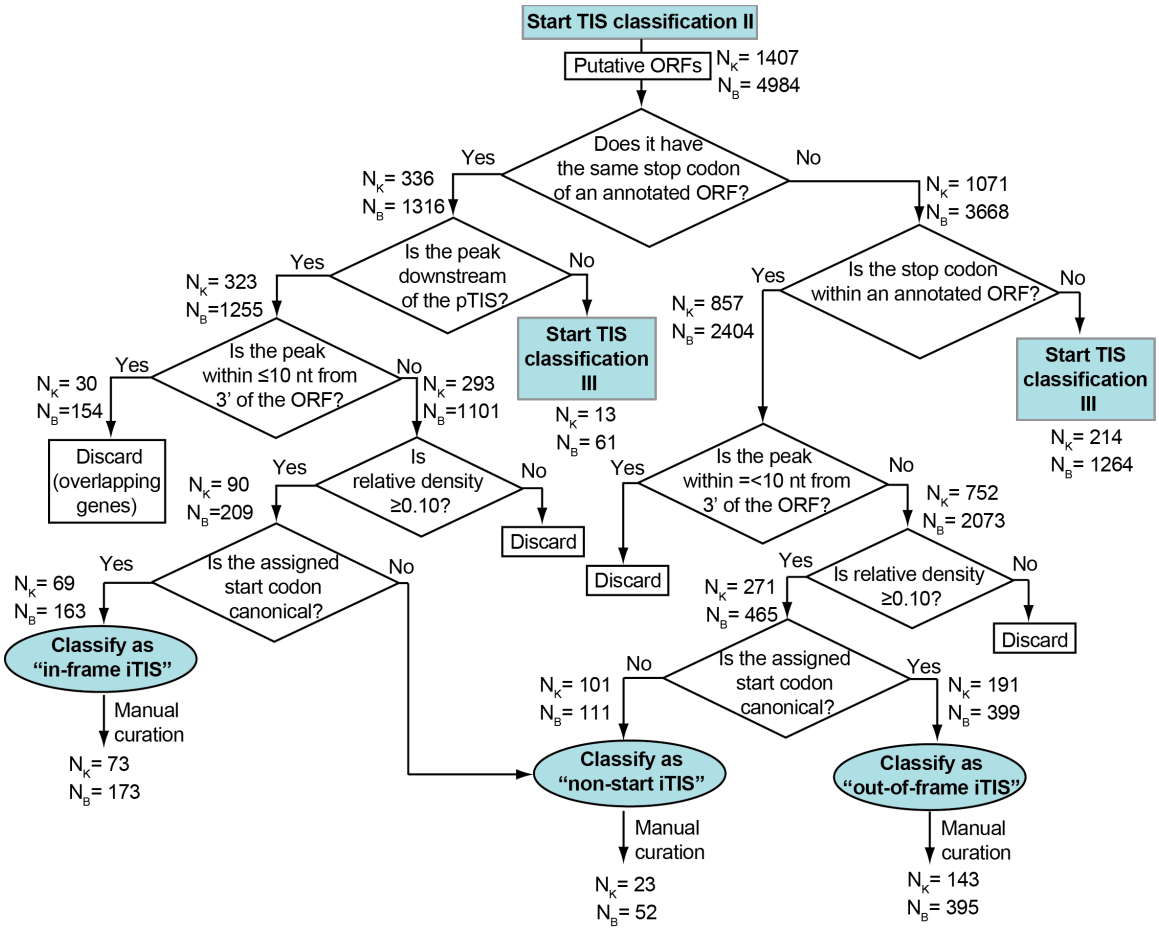
(A) Ribo-RET profile of the *potB* gene, shows no peak of the ribosome density at the start codon of the annotated pTIS (dark green flag), but instead reveals a strong peak at an in-frame start codon 27 nt upstream (pale green flag).

(B) In the *yjfN* gene, Ribo-RET reveals peak at the annotated pTIS (dark green flag) and an additional peak 9 nts upstream from it (marked with a pale green flag). The sequences surrounding the two TISs, including the SD-like regions (underlines) are shown.

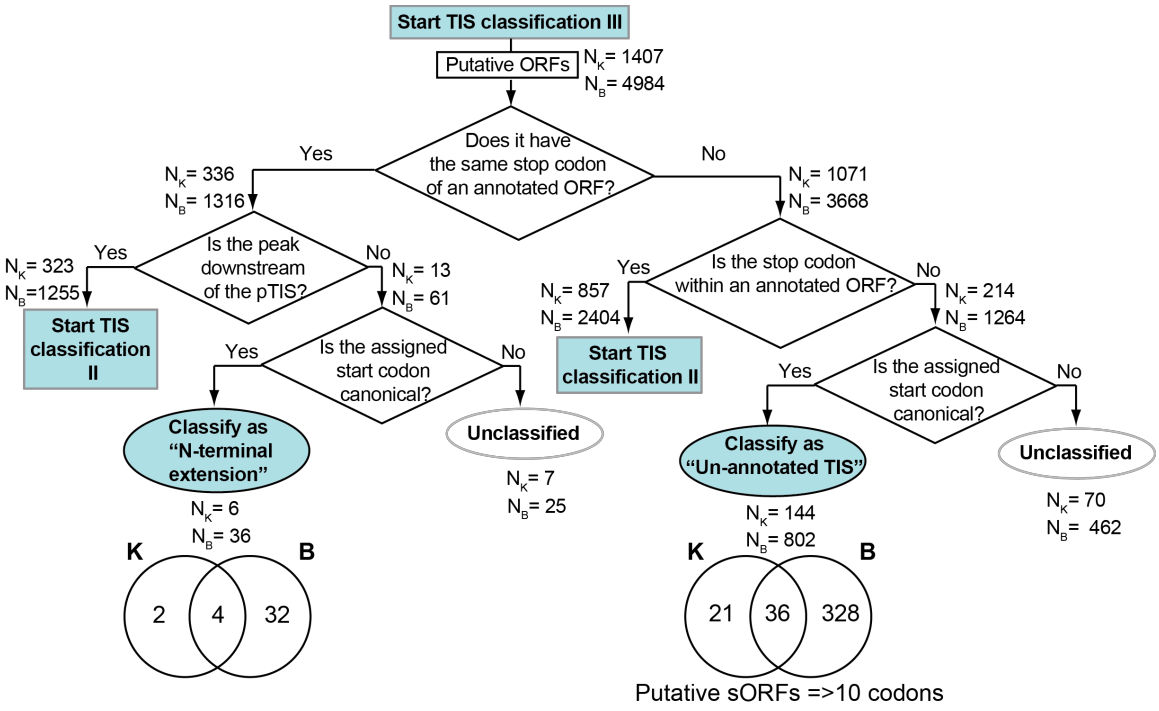
See Table S2 for other cases of Ribo-RET signals outside of the coding regions.



Supplemental scheme 1: Flow chart showing the criteria used to identify pTISs (Classification I), Related to STAR methods



Supplemental scheme 2: Flow chart showing the criteria used to identify iTISs (Classification II), Related to STAR methods



Supplemental scheme 3: Flow chart showing the criteria used to identify N-terminal extensions and un-annotated TISs (Classification III), Related to STAR methods

Table S3: List of primers and synthetic DNA fragments used in this study, Related to STAR methods

	Sequence (5' to 3')	Purpose
P1	TGTCCTGGCACTAATAGTGA	Forward primer for amplification of *tol::kan* cassette
P2	ACGATGCGTGGCGTATGG	Reverse primer for amplification of *tol::kan* cassette
P3	TTGTGAGCGGATAACAATTTCACACAGG AAACAGACCATGGTGGGTATTATTGGGG CAGG	Forward primer for amplification of *arcB*-PCR 1-wt
P4	ACATAATACTGCGCCAGC	Reverse primer for amplification of *arcB*-PCR 1-wt
P5	AAGCAAATTCGTCTGCTGG	Forward primer for amplification of *arcB*-PCR 2-wt
P6	TGGGAATATCGAGCAATGCTT	Reverse primer for amplification of *arcB*-PCR 2-wt
P7	GAAGAGAACAGTAAATCAGAAGCATTG	Forward primer for amplification of *arcB*-PCR 3
P8	TCAGGCTGAAAATCTTCTCTCATCCGCC AAAACAGCCAAGCTTTCACTTGTCATCGT CAT	Reverse primer for amplification of *arcB*-PCR 3
P9	TCCTGGGTATCCCAGAATTTC	Reverse primer for amplification of *arcB*-PCR 1-mutant
P10	CGCTAACCGCGATGATCAAGAAATTCTG GGATACCCAGGATGATGAAGAAAGTACG GTCACGACAGAAGAG	Forward primer for amplification of *arcB*-PCR 2-mutant
P11	TGGGAATATCGAGCAATGCTTCTGATTTA CTGTTCTCTTCTGTCGTGACCGTACTTTC TTCATCATCCTGGGTATCCCA	Reverse primer for amplification of *arcB*-PCR 2-mutant
#12	AACAGACCATGGTACCCAGGATGATGAG GAGAGTACGGTGACGACAGAAGAGAAC AGTAAATCAGAAGCATTGCTCGATATTCC CATGCTGGAACAGTATCTCGAACTTGTA GGACCGAAGCTGATCACCGACGGGTTA GCGGTGTTTGAGAAGATGATGCCGGGCT ATGTCAGCGTGCTGGAGTCGAATCTGAC GGCGCAGGATAAAAAAGGCATTGTTGAG GAAGGACATAAAATTAAAGGTGCGGCGG GGTCAGTGGGGTTACGCCATCTGCAACA GCTGGGTCAGCAAATTCAGTCTCCTGAC CTTCCGGCCTGGGAAGATAACGTCGGTG AATGGATTGAAGAGATGAAAGAAGAGTG GCGTCACGACGTAGAAGTGCTGAAAGC GTGGGTGGCAAAAGCCACTAAAAAAGAC TACAAAGACCATGACGGTGATTATAAAG ATCATGACATCGATTACAAGGATGACGA TGACAAGTGAAAGCTTGGCTGTT	gBlock for *arcB*-marker insert

P13	TAATACGACTCACTATAGGGCTGTAATTAACAACAAAGGGT	Forward primer for amplification of *atpB*
P14	GGTTATAATGAATTTTGCTTATTAACCGAGAATGTACGCAGTTAGTCCAGCTGAAGGTT	Reverse primer for amplification of *atpB*
P15	TAATACGACTCACTATAGGGACTAAAAGTAAGGCATTAAC	Forward primer for amplification of *mqo*
P16	GGTTATAATGAATTTTGCTTATTAACCTGCTCCTCGGACGCTTATTTCGCTTTTGCCGCC	Reverse primer for amplification of *mqo*
P17	GGTTATAATGAATTTTGCTTATTAAC	Reverse primer for toeprinting of *atpB* and *mqo*
P18	TAATACGACTCACTATAGGGAGCGCAGTGGAGACA	Forward primer for amplification of *birA*
P19	CTACGCAAATAATTTGCAGGG	Reverse primer for amplification of *birA*
P20	TTTCACCCAACTGCTC	Reverse primer for toeprinting of primary site of *birA*
P21	AATACTCCCCTTTCTTATTTTT	Reverse primer for toeprinting of internal site of *birA*
P22	TAATACGACTCACTATAGGGCAATAACAAGGATTGTCGCAATG	Forward primer for amplification of *sfsA*-PCR 1
P23	GCCGTATTTTACTTCGCTTTCTAGCGAGCTATAGCCTGACAG	Reverse primer for amplification of *sfsA*-PCR 1
P24	CTGTCAGGCTATAGCTCGCTAGAAAGCGAAGTAAAATACGGC	Forward primer for amplification of *sfsA*-PCR 2
P25	CTACAATGTAACCGGCAGTG	Reverse primer for amplification of *sfsA*-PCR 2
P26	TAATACGACTCACTATAGGG	Forward primer for amplification of *sfsA*-g321a, a322g
P27	CTACAATGTAACCGGCAGTG	Reverse primer for amplification of *sfsA*-g321a, a322g
P28	CATCGGGTGTGATCAC	Reverse primer for toeprinting of primary site of *sfsA*
P29	CGATTTCACTTCAATATA	Reverse primer for toeprinting of internal site of *sfsA*
P30	TAATACGACTCACTATAGGGGCTAATGTGAAGGAGACGC	Forward primer for amplification of *hslR*
P31	TTATTCACTGTCGCCGTG	Reverse primer for amplification of *hslR*
P32	GGGCCAGCGCGC	Reverse primer for toeprinting of primary site of *hslR*
P33	TTGTCCGGGCGTCGG	Reverse primer for toeprinting of internal site of *hslR*
P34	TAATACGACTCACTATAGGGAATGCTATCAGGAGTTTACGATG	Forward primer for amplification of *yecJ*
P35	TTAATGGGATTCACCCTGTGGG	Reverse primer for amplification of *yecJ*
P36	CATCCAGAATTTGTTTGATAAC	Reverse primer for toeprinting of primary site of *yecJ*

P37	GCGGCGGCGGGATGG	Reverse primer for toeprinting of internal site of *yecJ*
P38	GTGAGCGGATAACAATTTCACACAGAAT TCATTAAAGAGGAGAAATTAACTATGGCT	Forward primer for amplification of *RFP*
P39	ATATCTCCTTCTTAAAGTTAAACAACTAG TCTATTCGCCAGAACCAGC	Reverse primer for amplification of *RFP*
P40	TTTAAGAAGGAGATATACATATGACTAGT GCATCCAAGGGCGA	Forward primer for amplification of *GFP*
P41	TCAGCTAATTAAGCTTGGCTGCAGGTCG ACCCGGGGTACCGAG	Reverse primer for amplification of *GFP*
P42	TCCGCTGCTGGTTCTGGCGAATAGACTA GTCAATAACAAGGATTGTCGCAATG	Forward primer for amplification of insert for pRXGSM-sfsA-wt
P43	AAAGAGCTCCTCGCCCTTGGATGCACTA GTGCGAGCTATAGCCTGAC	Reverse primer for amplification of insert for pRXGSM-sfsA-wt
P44	AAAGAGCTCCTCGCCCTTGGATGCACTA GTGCGAGCTATAGCCTGACAGTTCTGAA ATTGATTCGATAAGGATAGCCT	Reverse primer for amplification of insert for pRXGSM-sfsA-mutant
P45	GTTCTGGCGAATAGACTAGTAAATGCTAT CAGGAGTTTACG	Forward primer for amplification of insert for pRXGSM-yecJ derivatives
P46	AGCTCCTCGCCCTTGGATGCACTAGTCA TCGAGAACATCCAGAATTTG	Reverse primer for amplification of insert for pRXGSM-yecJ-iTIS-wt
P47	AGCTCCTCGCCCTTGGATGCACTAGTCG TCGAGAACATCCAGAATTTG	Reverse primer for amplification of insert for pRXGSM-yecJ-iTIS(-)
P48	AGAGCTCCTCGCCCTTGGATGCACTAGT TACATCGAGAACATCCAGAATTTG	Reverse primer for amplification of insert for pRXGSM-yecJ-pTIS-wt
P49	CTGCTGGTTCTGGCGAATAGACTAGTAA TGCTATCAAAAGTTTACGTCGTCCCAGC CGCT	Primer for site directed mutagenesis to generate pRXGSM-yecJ-pTIS(-)
P50	AGCTCCTCGCCCTTGGATGCACTAGTTA CGTCGAGAACATCCAGAATTTG	Reverse primer for amplification of insert for pRXGSM-yecJ-pTIS-iTIS(-)
P51	AGCTCCTCGCCCTTGGATGCACTAGTGA CATCGAGAACATCCAGAATTTG	Primer for site directed mutagenesis to generate pRXGSM-yecJ-pTIS-iStop(-)
P52	GGCCTTAACCGCTAACGT	Direct primer for sequencing the iTIS region in the arcB gene
P53	TTTAATCTGTATCAGGCTGAAAATCTT	Reverse primer for sequencing the iTIS region in the *arcB* gene

Supplementary Information References

Arenz, S., Bock, L.V., Graf, M., Innis, C.A., Beckmann, R., Grubmüller, H., Vaiana, A.C., and Wilson, D.N. (2016). A combined cryo-EM and molecular dynamics approach reveals the mechanism of ErmBL-mediated translation arrest. Nat. Commun. *7*, 12026.

Baek, J., Lee, J., Yoon, K., and Lee, H. (2017). Identification of unannotated small genes in *Salmonella*. G3 (Bethesda) *7*, 983-989.

Bienvenut, W.V., Giglione, C., and Meinnel, T. (2015). Proteome-wide analysis of the amino terminal status of *Escherichia coli* proteins at the steady-state and upon deformylation inhibition. Proteomics *15*, 2503-2518.

Buch, J.K., and Boyle, S.M. (1985). Biosynthetic arginine decarboxylase in *Escherichia coli* is synthesized as a precursor and located in the cell envelope. J. Bacteriol. *163*, 522-527.

Davidovich, C., Bashan, A., Auerbach-Nevo, T., Yaggie, R.D., Gontarek, R.R., and Yonath, A. (2007). Induced-fit tightens pleuromutilins binding to ribosomes and remote interactions enable their selectivity. Proc. Natl. Acad. Sci. U. S. A. *104*, 4291-4296.

Kannan, K., Kanabar, P., Schryer, D., Florin, T., Oh, E., Bahroos, N., Tenson, T., Weissman, J.S., and Mankin, A.S. (2014). The general mode of translation inhibition by macrolide antibiotics. Proc. Natl. Acad. Sci. U. S. A. *111*, 15958-15963.

Li, G.W., Burkhardt, D., Gross, C., and Weissman, J.S. (2014). Quantifying absolute protein synthesis rates reveals principles underlying allocation of cellular resources. Cell *157*, 624-635.

Nakahigashi, K., Takai, Y., Kimura, M., Abe, N., Nakayashiki, T., Shiwa, Y., Yoshikawa, H., Wanner, B.L., Ishihama, Y., and Mori, H. (2016). Comprehensive identification of translation start sites by tetracycline-inhibited ribosome profiling. DNA Res. *23*, 193-201.

Nakatogawa, H., and Ito, K. (2002). The ribosomal exit tunnel functions as a discriminating gate. Cell *108*, 629-636.

Polikanov, Y.S., Steitz, T.A., and Innis, C.A. (2014). A proton wire to couple aminoacyl-tRNA accommodation and peptide-bond formation on the ribosome. Nat. Struct. Mol. Biol. *21*, 787-793.