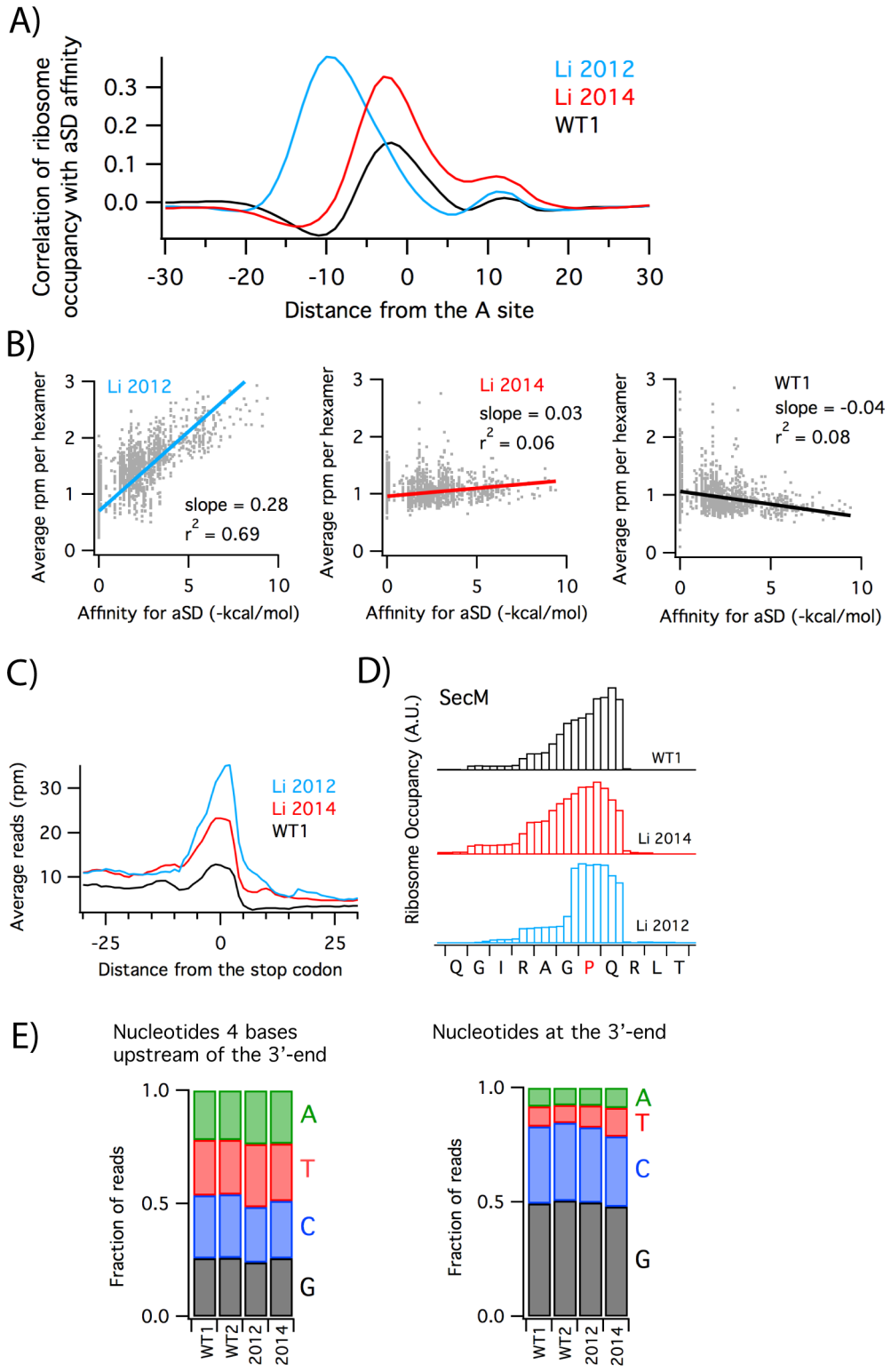
Table S1. Linear fits for aSD affinity and ribosome density correlations, related to Figure 2

Citation	slope	r^2 value	SRR run #	This study:	slope	r^2 value
Li et al. 2012	0.28	0.67	SRR407274-5	WT1	-0.01	0.01
Li et al. 2014	0.07	0.29	SRR1067765-8	WT2	-0.01	0.01
Oh et al. 2011	0.08	0.17	SRR364364	WT3	-0.01	0.00
Oh et al. 2011	0.06	0.12	SRR364366	WT4	-0.02	0.02
Oh et al. 2011	0.15	0.55	SRR364368	WT5	-0.03	0.04
Oh et al. 2011	0.13	0.46	SRR364370	WT6	-0.01	0.00
Balakrishnan et al. 2014	-0.10	0.23	SRR1613263	WT7	0.00	0.00
Balakrishnan et al. 2014	-0.09	0.19	SRR1613265	WT8	0.00	0.00
Balakrishnan et al. 2014	-0.09	0.20	SRR1613266	WT9	0.00	0.00
Elgamal et al. 2014	0.13	0.31	SRR1200750	WT10	0.01	0.01
Elgamal et al. 2014	0.08	0.18	SRR1200751	WT11	0.03	0.04
Haft et al. 2014	0.21	0.47	SRR1211047	WT12	-0.02	0.02
Haft et al. 2014	0.25	0.55	SRR1211048	WT13	-0.06	0.11
Kannan et al. 2014	0.09	0.25	SRR1583082	WT14	-0.01	0.00
Subramaniam et al. 2014	0.15	0.47	SRR1301057	WT15	-0.05	0.20
Subramaniam et al. 2014	0.05	0.09	SRR1301059	WT16	-0.03	0.04
Liu et al. 2013	0.08	0.18	SRR869826	WT17	-0.04	0.06
Liu et al. 2013	0.10	0.23	SRR869827	WT18	-0.02	0.04
Guo et al. 2014	0.03	0.06	SRR1425203	WT19	-0.03	0.05
Guo et al. 2014	0.01	0.00	SRR1425204			

In this table we show the parameters for the linear fits of aSD affinity and ribosome density for mRNA hexamers as depicted in Figure 2A. On the right side are 19 libraries created in this study; these reproducibly show no correlation between aSD affinity and ribosome density. In contrast, on the left side, 20 *E. coli* profiling libraries from other labs exhibit a high degree of variability. Although these libraries are nearly all from wild-type controls, they were prepared in very different ways. Some labs collect cells with filtering and others with centrifugation; some add chloramphenicol to the media and others do not; some lyse by grinding cell pellets and others by freeze / thaw cycles. Particularly relevant to our discussion is the fact that these studies all report isolating 28 – 42 nt fragments during the size selection step, following the Weissman protocol, except for the Balakrishnan et al. 2014 study, which selected 20 – 30 nt fragments. The negative correlation in Balakrishnan samples comes from

preferentially cloning shorter fragments that lack SD motifs—the opposite of the enrichment in the Li et al. 2012 study.

Figure S1: SD pauses as observed in center-assigned density, related to Figure 1



We demonstrated elsewhere that assigning ribosome occupancy to the 3'-end of reads in bacterial profiling data yields a more precise and accurate view of the position of the ribosome (Woolstenhulme et al., 2015). We note, however, that SD pauses were initially observed using center-assignment and it is possible that 3'-assignment interferes with our ability to accurately detect SD pauses. Indeed, we see in Figure 1A that 3'-assignment of the Li et al. 2012 data reduces the SD pausing signal in cross-correlation plots. This is because SD motifs tend to be near the 5'-end of reads, where they vary in distance from the 3'-end according to the caterpillar model of O'Connor et al. (2013). In spite of this weaker signal, we find that our linear fits of the data from Li et al. 2012 using 3'-assignment reproduce the strong correlation that was previously reported (Figure 2A). In addition, we show here using center-assignment that the strong pausing signal reported in Li et al. 2012 is lacking in the 2014 data and in our libraries.

The 2012 data have a maximal correlation between aSD affinity and ribosome density 8 – 11 nt upstream of the A site codon (Figure S1A). In contrast, the highest correlation in the 2014 and WT1 data occurs 3 nt upstream of the A site codon. Given that biochemical and structural studies show that the A site codon is 12 nt upstream of the 3'-boundary of the ribosome, these distances correspond exactly with the –22 and –15 peaks observed in Figure 1A using 3'-assignment. We argue that given their different positions, these two peaks are fundamentally distinct: the peak in the 2012 data arises from a true SD correlation whereas the peak in the 2014 and WT1 data arises predominantly from pauses on Gly codons.

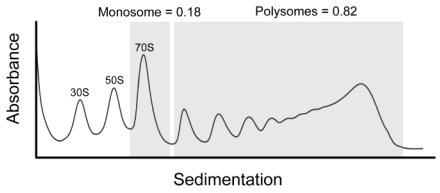
Using the center-assignment strategy, we calculated the average ribosome density and aSD affinity for all RNA hexamers. The average ribosome density was computed for the region 11 – 17 nt downstream of the first nt in the hexamer. As shown in Figure S1B, we see a strong correlation in the 2012 data, with the same slope and a similar r^2 value as was observed in Figure 2A using 3'-assignment. In contrast, little or no correlation is seen for the 2014 and WT1 data. These findings show that our conclusions are equivalent whether ribosome occupancy is assigned using the center or 3'-end strategies.

A clear weakness of the center-assignment strategy is that the ribosome density maps need to be shifted manually in order to line them up with the ribosomal A site. In the 2012

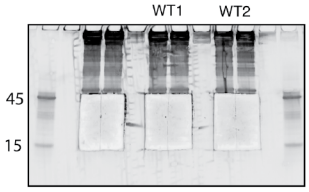
paper, the Li et al. shifted the density maps 4 nt downstream so that the observed density lines up with the A site at stop codons and known translational stalling sites like SecM. The shift depends on the length distribution of mRNA fragments in the library and has to be determined empirically for each individual library. Here we provide three pieces of evidence that our center-assigned density maps are shifted properly. We found that the 2014 and WT1 maps did not require shifting because they naturally line up with the A site at stop codons (Figure S1C) and the well-characterized arrest at SecM in which the Pro codon is positioned in the A site (Figure S1D). It makes sense that density maps from these libraries do not require shifting: their mRNA fragments are shorter than the 2012 library. Given that the RNA fragments differ in length almost exclusively at the 5'-ends, fragments in these libraries do not have extra 5'-sequence that pulls the distribution upstream. Finally, we note that there are small peaks in the cross-correlation plots 12 nt downstream of the A site (Figure S1A). These peaks arise from cloning bias at the 3'-end of the RNA fragments. Although neutral positions with the fragments show no enrichment for specific nucleotides (Figure S1E, left panel), the 3'-end of cloned fragments is enriched in G (Figure S1E, right panel). Since G-rich sequences have high aSD affinity, this creates a peak in the cross-correlation plot. Given that the distance between the 3'-end and the A site is known to be constant in ribosome profiling reads (Woolstenhulme et al., 2015), the fact that these small peaks line up in all three libraries in Figure S1A means that the density maps are shifted correctly and consistently. Proper alignment of the density maps is essential to calculating pausing at the same position in the ribosome across different libraries using center-assignment.

Figure S2. Quality control metrics for our libraries, related to Figure 1 and 2

A)

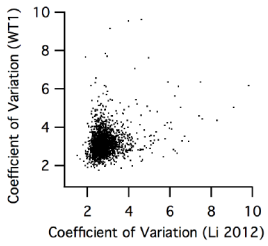


B)



RNA size selection gel, post-excision

C)



D)

A-site codon	WT1	WT2	WT3	Li 2012	Li 2014
A	0.95	0.96	0.99	0.70	0.75
C	1.05	1.06	0.92	1.53	1.52
E	1.38	1.30	1.79	1.27	1.38
D	1.46	1.42	1.95	1.36	1.28
G	1.08	1.20	0.74	0.83	0.90
F	1.09	1.03	1.20	2.01	1.86
I	1.05	1.01	1.22	1.23	0.99
H	1.24	1.13	1.58	1.05	0.90
K	1.44	1.29	1.84	1.39	1.29
M	1.02	0.96	1.13	1.39	1.25
L	1.09	1.02	1.21	1.58	1.54
N	1.07	1.01	1.38	1.19	1.01
Q	1.21	1.16	1.48	1.39	1.27
P	1.25	1.24	1.22	1.33	1.61
S	2.46	2.64	0.94	1.40	1.38
R	1.51	1.38	1.97	1.32	1.32
T	2.04	2.02	1.39	1.75	1.72
W	1.28	1.25	1.47	1.28	1.40
V	1.28	1.23	1.40	1.58	1.54
Y	1.23	1.20	1.66	1.61	1.21

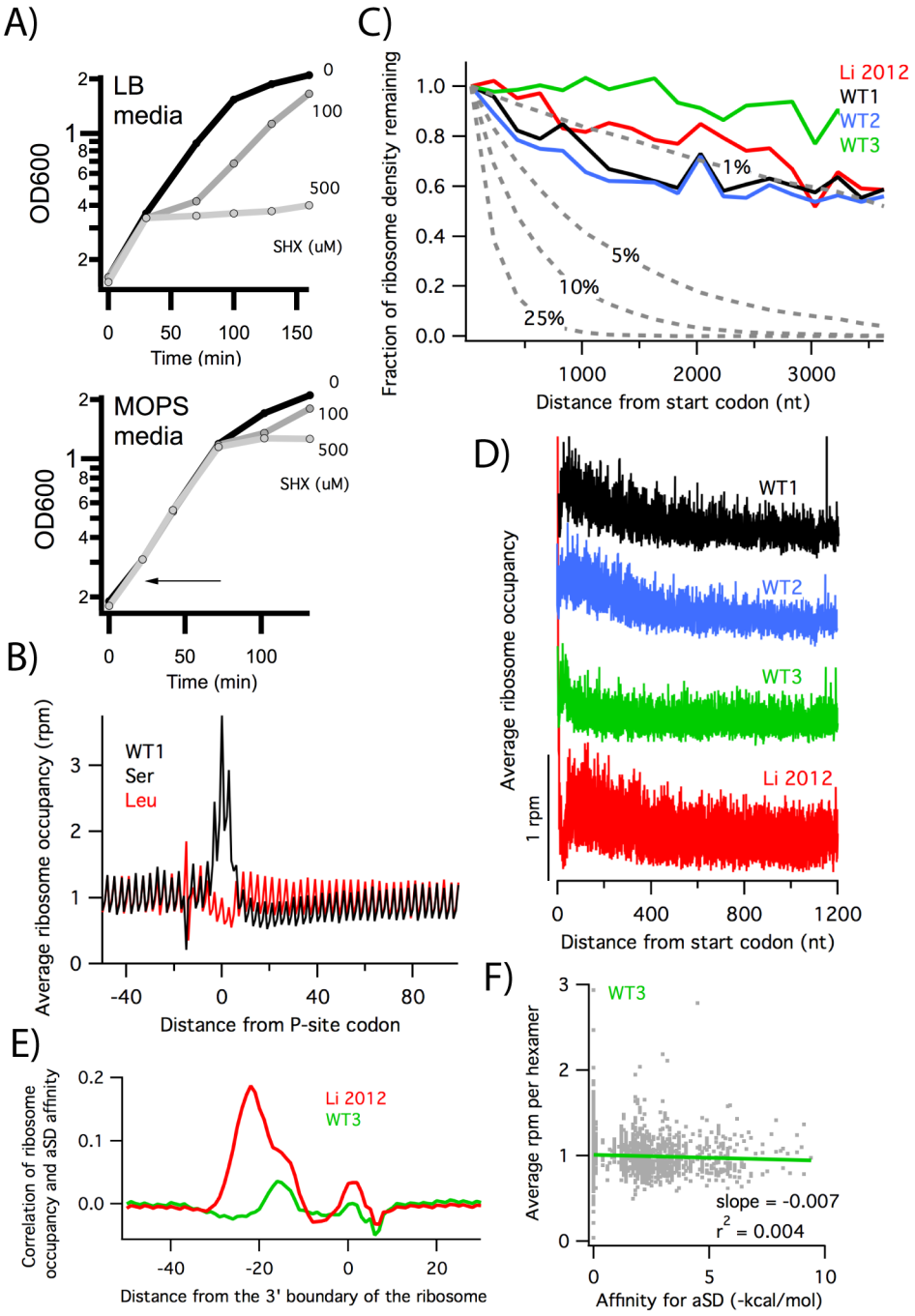
We worked very hard to replicate the SD pauses observed by Li et al. 2012 using the same strain (MG1655) and growth conditions. To match their protocol, we monitored growth rates and optimized the media formulation, titrated the MNase concentration against an aliquot from the Weissman lab to ensure we were using similar levels of enzyme activity, and experimented with variations in the filtering and freezing protocols. As we argue in the main body of the text, the clear difference in our protocols is the purification of mRNA fragments: we sampled the distribution of fragments broadly (15 – 45 nt) whereas Li et al. 2012 only sampled longer reads (28 – 42 nt). But we also wanted to rule out any confounding differences in the protocol or problems in our library construction. Despite all of our efforts, we have been unable to detect significant correlations between SD motifs and ribosome density. Here we discuss quality control metrics for steps that could conceivably be important to observe pausing in ribosome profiling data.

First, we see robust polysome signals in sucrose gradients (Figure S2A) showing that the cells are healthy and have high levels of translation at the point of harvesting. The profile also indicates that we have not lost ribosomes or mRNA integrity during the lysis process. After digestion with MNase, we recovered ribosomes quantitatively as the polysome fraction collapsed into monosomes, indicating that we are not losing mRNA fragments or biasing the library at this step (not shown). In our size selection gels, fragments 15 – 45 nt in length were isolated using RNA markers as size controls (Figure S2B). This gel and the read length distributions in Figure 1B show that we captured the relevant ribosome protected fragments and did not lose SD-containing reads by selecting only shorter reads.

It is conceivable that SD pauses occur in our data but we cannot see them because they are masked by noise of greater intensity. We computed the coefficient of variation for genes with more than one read per codon on average in our WT1 library and the data from Li et al. 2012. The coefficient of variation gives a rough idea of the variability in the ribosome density across each gene. As is clear from the values for many genes that are plotted in Figure S2C, the noise does not differ greatly between the two datasets, refuting the suggestion that overall noise in our data prevented us from observing SD pauses.

Moving from the general to the specific, we calculated pause scores for codons in the ribosomal A site grouped by the encoded amino acid (Figure S2D)—an estimate of pauses that occur as the ribosome waits for incoming aminoacyl-tRNA. These were calculated by dividing the density at the first nt of the A site codon by the mean for the entire gene and averaging the scores for every relevant codon throughout the genome. We find that density at Ser and Thr codons is elevated in our WT1 and WT2 samples compared to other amino acids. Although Thr seems to also be high in the libraries of Li et al., Ser pauses appear to be a difference in our samples. Potential implications are discussed below in Figure S3. Importantly, however, we note that the range of pause scores is not dramatically different in our WT1 (0.95 to 2.46) and the 2012 data (0.7 to 2.0).

Figure S3. Ser pauses do not explain the lack of SD pauses in our libraries, related to Figure 1 and 2



As noted in Figure S2D, pauses at Ser codons are higher in our WT1 and WT2 libraries than in the Li et al. 2012 library. We wondered whether this pausing signal might lead to loss of ribosome density at the 3'-end of genes as ribosomes are removed from the message by rescue mechanisms such as the tmRNA pathway (Subramaniam et al., 2014). If substantial loss of ribosome density occurs along genes, calculations of SD pause strengths would be inaccurate because the signal would vary depending on the position of the motif within the gene.

Because we suspected that Ser pauses arose from starvation due to problems with the media formulation or growth conditions, we harvested cells in early log phase in a complete synthetic MOPS medium with high concentrations of Ser and glucose. It is well documented that upon depletion of glucose in LB at around OD_{600} = 0.3, Ser becomes limiting for translation as it is metabolized (Li et al., 2012; Pruss et al., 1994; Sezonov et al., 2007). We confirmed this by adding 500 μM serine hydroxamate (SHX) to LB media and observing an arrest of cell growth (Figure S3A). In contrast, adding SHX to our MOPS media had no effect on growth until stationary phase, indicating that there was abundant Ser present when our culture was harvested in early log phase (OD_{600} = 0.25, indicated with an arrow). We conclude that the Ser pauses are not the result of starvation during the growth of the culture.

We also observed that Ser pauses in our data have an unexpected effect on the ribosome density downstream (Figure S3B). Although there is a reduction of about 25% immediately after Ser codons, ribosome density recovers to its original level by 80 nt downstream. The fact that the pause is only locally rate-limiting suggests that we are observing a time-dependent event, similar to the time-dependent run-off of ribosomes that occurs if harringtonine is added to trap ribosomes at start codons (Ingolia et al., 2011). The dip in density downstream of Ser pauses is consistent with continued elongation lengthening the distance between the paused ribosome and downstream ribosomes. This may be evidence of translation (and pausing) in the lysate.

Another way to measure the decay of ribosome density along genes is to compute the fraction of the density remaining in 200 nt windows compared to the density at the 5'-end of the gene. Although we do not believe that Ser pauses are a strong contributor, we do find that

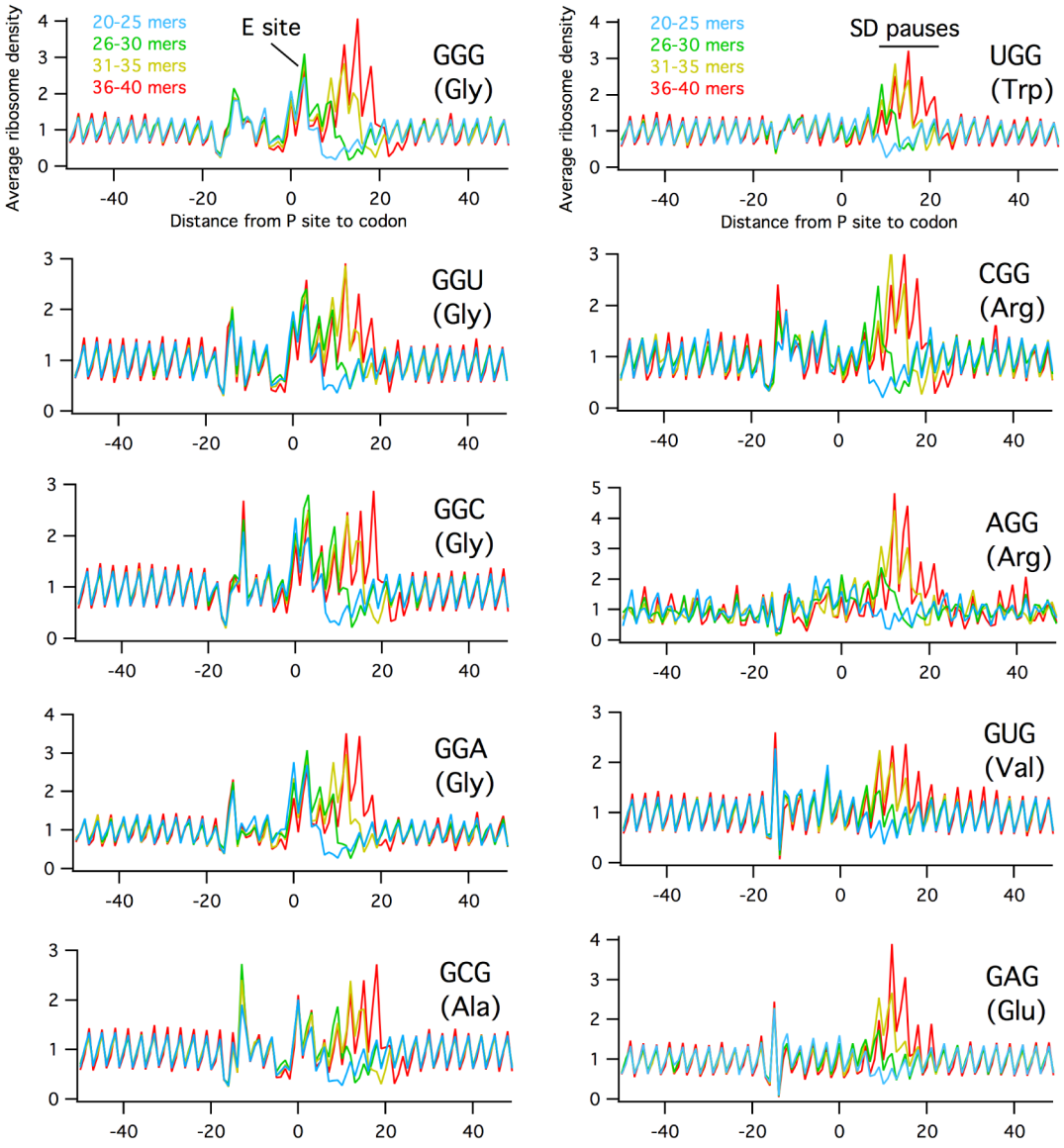
there is less ribosome density at the 3'-end of genes in our data compared with the 5'-end (Figure S3C). For our WT1 and WT2 samples, we observe a 40% reduction of density by about 1200 nt after the start codon. In comparison, there is a 20% reduction in density in the Li et al. 2012 data at this position. These plots are not consistent with a loss of ribosomes from messages after Ser codons. Simulations using the same set of genes reveal exponential decay as expected; even a 5% loss of ribosomes after Ser codons leads to a far more rapid decay of density than observed in our data (dotted lines, Figure S3C). The simulations are perhaps consistent with a 1% loss of ribosomes after Ser codons in the WT1 library, but the shape of the plots provides additional clues that suggest another origin.

For WT1 and WT2, the density drops early in the gene and remains fairly constant at a plateau thereafter (Figure S3C). Given that the decay curves are calculated by dividing the downstream density by the density near the start codon, this could be explained by higher density at the 5'-end of genes as observed in other profiling studies (Ingolia et al., 2009). The same phenomenon can also be seen in plots of average density that include genes longer than 1200 nt aligned at the start codon (Figure S3D). The 5'-ramp probably arises from continuing initiation in the presence of imperfect elongation inhibitors (Gerashchenko and Gladyshev, 2014). We argue that it is more likely that ribosomes continue to be loaded at the 5'-end during the preparation of the samples (perhaps during filtering or freezing) than it is that they are being lost from messages at strong pause sites.

Most importantly, our WT3 library has no detectable SD pauses despite the fact that it doesn't have these confounding factors. This library was prepared by a different procedure in which we filtered the cells completely dry prior to freezing them in liquid nitrogen, following the Weissman lab protocol, rather than scraping cells off of the filter before the media runs dry (our usual protocol). We found that unlike the WT1 library, WT3 has no Ser pauses (Figure S2D) and no pauses that result in even a temporary loss of density downstream. We also observed that the WT3 data showed very little decay in ribosome density, even less than the Li 2012 data (Figure S3C). The lack of apparent decay is probably due to the fact that there is essentially no 5'-ramp (Figure S3D). This may indicate that with this cell harvesting protocol, there is less translation in the lysate. Importantly, in this sample, even without these confounding factors,

we do not observe pauses at SD motifs, whether at the –22 position in cross-correlation plots or in the linear fit of SD affinity and ribosome density (Figures S3E and S3F). These findings together refute the suggestion that we missed SD pausing due to quality issues with our data (either from Ser pauses or loss of density along genes).

Figure S4. Gly and SD pauses on all ten G-rich codons, related to Figure 4.



These plots show average ribosome density at all ten codons containing two G nucleotides using density maps made with various read lengths as in Figure 4. The peaks at 15 – 20 are consistent with SD pauses both in their position and their read length dependence. The peak at 3 corresponds to pauses in the ribosomal E site. These pauses are stronger in Gly codons (GGN) than the other six as quantitated in Figure 4C. The peaks between –15 and –20 arise from cloning bias at the 3'-end of reads (see Figure S1E and the peak at 0 in Figure 1A).

Supplemental Experimental Procedures

In vitro translation constructs:

All toeprinting DNA templates start with the following 5′ sequence that includes a T7 promoter, ribosome binding site, and start codon (underlined): CTGTACATTAATACGACTCACTATAGGGAGATTTTATAAGGAGGAAAAAATATG. The 3′ end of all templates includes a binding site for the NV1 primer, GGTTATAATGAATTTTGCTTATTAAC. To characterize internal Shine-Dalgarno sequences, four sites from endogenous *E. coli* genes were chosen, GGUGGU in both *ompF* (at 420) and *atpA* (at 1131), GGAGGU in *cyoB* (at 78), and AGGAGG in *mliC* (at 151). In our constructs, 33 nt of the natural sequence was inserted in the correct reading frame after a constant upstream region, with the SD-motif starting at position 13 in the 33 nt sequence. The final DNA constructs were as follows:

From *ompF* encoding MISVNGALPEFGGDTAYSIA-stop:
CTGTACATTAATACGACTCACTATAGGGAGATTTTATAAGGAGGAAAAAATATGATTTCCGTGAACGGC
GCACTGCCAGAATTTGGTGGTGATACTGCATACAGCATTGCCTAAGTAAGTAAAGATCTTAGGCGCGCC
GGATCTGCATCGTTAATAAGCAAAATTCATTATAACC

From *atpA* encoding MISVNGAVSRVGGAAQTKIA-stop:
CTGTACATTAATACGACTCACTATAGGGAGATTTTATAAGGAGGAAAAAATATGATTTCCGTGAACGGC
GCAGTATCCCGTGTTGGTGGTGCAGCACAGACCAAGATTGCCTAAGTAAGTAAAGATCTTAGGCGCGCC
GGATCTGCATCGTTAATAAGCAAAATTCATTATAACC

From *cyoB* encoding MISVNGAGIILGGLALVGIA-stop:
CTGTACATTAATACGACTCACTATAGGGAGATTTTATAAGGAGGAAAAAATATGATTTCCGTGAACGGC
GCAGGCATTATTTTGGGAGGTCTGGCGCTCGTTGGCATTGCCTAAGTAAGTAAAGATCTTAGGCGCGCC
GGATCTGCATCGTTAATAAGCAAAATTCATTATAACC

From *mliC* encoding MISVNGANPRQEVSFVYDIA-stop:

CTGTACATTAATACGACTCACTATAGGGAGATTTTATAAGGAGGAAAAAATATGATTTCCGTGAACGGC
GCAAATCCGCGCCAGGAGGTCAGTTTTGTTTACGATATTGCCTAAGTAAGTAAAGATCTTAGGCGCGCC
GGATCTGCATCGTTAATAAGCAAAATTCATTATAACC

As a control, we also considered pausing at the Pro-Pro-Met motif at position 507 in the *gltJ* gene. 39 nt of the natural sequence were inserted in the proper reading frame after the constant region, with the Pro-Pro-Met motif starting at position 24 of the 39 nt sequence.

From *gltJ* encoding MISVNGAPNAYRVIVPPMTSIA-stop:
CTGTACATTAATACGACTCACTATAGGGAGATTTTATAAGGAGGAAAAAATATGATTTCCGTGAACGGC
GCACCTAATGCTTATCGCGTTATCGTCCCGCCGATGACCTCAATTGCCTAAGTAAGTAAAGATCTTAGGC
GCGCCGGATCTGCATCGTTAATAAGCAAAATTCATTATAACC

Toeprinting analyses:

The PURExpress system (New England Biolabs) was used for *in vitro* translation. 0.2 pmol of template DNA was combined on ice with 2 µl of Solution A and 1.5 µl of Solution B along with either 0.5 µl water or thiostrepton (0.5 mm in 5% DMSO), then incubated at 37 °C for 30 min. 1 pmol of [^{32}P]ATP-labeled NV1 primer was added to each reaction along with 2 U of Ambion SUPERasin RNase Inhibitor (Life Technologies). After incubation at 37 °C for 2 min, the samples were placed on ice for 5 min and at 25 °C for 5 min. Reverse transcription was performed by supplementing each sample with a mixture of four dNTPs to a final concentration of 0.32 mM each, adding 2.4 U of AMV Reverse Transcriptase (Roche), and incubating at 37 °C for 15 min. Reactions were stopped and the RNA hydrolyzed by addition of 1 µl 10 N NaOH, incubation at 37 °C for 15 min, and neutralization with 0.8 µl 12 M HCl. Samples were then diluted with 200 µl of extraction buffer (0.3 M Na-acetate, 0.5% SDS, 5 mM EDTA pH 8.0) and extracted with phenol and chloroform. After ethanol precipitation, pellets were resuspended in 6 ul of formamide-EDTA loading dye (90% formamide, 25 mM EDTA, pH 8.0) and separated by 8% denaturing PAGE and visualized with a Typhoon FLA 9500 (GE).

Supplemental References

Balakrishnan, R., Oman, K., Shoji, S., Bundschuh, R., and Fredrick, K. (2014). The conserved GTPase LepA contributes mainly to translation initiation in Escherichia coli. Nucleic Acids Res *42*, 13370-13383.

Elgamal, S., Katz, A., Hersch, S.J., Newsom, D., White, P., Navarre, W.W., and Ibba, M. (2014). EF-P dependent pauses integrate proximal and distal signals during translation. PLoS Genet *10*, e1004553.

Gerashchenko, M.V., and Gladyshev, V.N. (2014). Translation inhibitors cause abnormalities in ribosome profiling experiments. Nucleic Acids Res *42*, e134.

Guo, M.S., Updegrove, T.B., Gogol, E.B., Shabalina, S.A., Gross, C.A., and Storz, G. (2014). MicL, a new sigmaE-dependent sRNA, combats envelope stress by repressing synthesis of Lpp, the major outer membrane lipoprotein. Genes & development *28*, 1620-1634.

Haft, R.J., Keating, D.H., Schwaegler, T., Schwalbach, M.S., Vinokur, J., Tremaine, M., Peters, J.M., Kotlajich, M.V., Pohlmann, E.L., Ong, I.M.*, et al.* (2014). Correcting direct effects of ethanol on translation and transcription machinery confers ethanol tolerance in bacteria. Proc Natl Acad Sci U S A *111*, E2576-2585.

Ingolia, N.T., Ghaemmaghami, S., Newman, J.R., and Weissman, J.S. (2009). Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. Science *324*, 218-223.

Ingolia, N.T., Lareau, L.F., and Weissman, J.S. (2011). Ribosome profiling of mouse embryonic stem cells reveals the complexity and dynamics of mammalian proteomes. Cell *147*, 789-802.

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.

Li, G.W., Oh, E., and Weissman, J.S. (2012). The anti-Shine-Dalgarno sequence drives translational pausing and codon choice in bacteria. Nature *484*, 538-541.

Liu, X., Jiang, H., Gu, Z., and Roberts, J.W. (2013). High-resolution view of bacteriophage lambda gene expression by ribosome profiling. Proc Natl Acad Sci U S A *110*, 11928-11933.

Oh, E., Becker, A.H., Sandikci, A., Huber, D., Chaba, R., Gloge, F., Nichols, R.J., Typas, A., Gross, C.A., Kramer, G.*, et al.* (2011). Selective ribosome profiling reveals the cotranslational chaperone action of trigger factor in vivo. Cell *147*, 1295-1308.

Pruss, B.M., Nelms, J.M., Park, C., and Wolfe, A.J. (1994). Mutations in NADH:ubiquinone oxidoreductase of Escherichia coli affect growth on mixed amino acids. J Bacteriol *176*, 2143-2150.

Sezonov, G., Joseleau-Petit, D., and D'Ari, R. (2007). Escherichia coli physiology in Luria-Bertani broth. J Bacteriol *189*, 8746-8749.

Subramaniam, A.R., Zid, B.M., and O'Shea, E.K. (2014). An integrated approach reveals regulatory controls on bacterial translation elongation. Cell *159*, 1200-1211.

Woolstenhulme, C.J., Guydosh, N.R., Green, R., and Buskirk, A.R. (2015). High-precision analysis of translational pausing by ribosome profiling in bacteria lacking EFP. Cell reports *11*, 13-21.