LETTER

doi:10.1038/nature10965

# The anti–Shine–Dalgarno sequence drives translational pausing and codon choice in bacteria

Gene-Wei Li[1], Eugene Oh[1] & Jonathan S. Weissman[1]

**Protein synthesis by ribosomes takes place on a linear substrate but at non-uniform speeds. Transient pausing of ribosomes can affect a variety of co-translational processes, including protein targeting and folding[1]. These pauses are influenced by the sequence of the messenger RNA[2]. Thus, redundancy in the genetic code allows the same protein to be translated at different rates. However, our knowledge of both the position and the mechanism of translational pausing *in vivo* is highly limited. Here we present a genome-wide analysis of translational pausing in bacteria by ribosome profiling—deep sequencing of ribosome-protected mRNA fragments[3–5]. This approach enables the high-resolution measurement of ribosome density profiles along most transcripts at unperturbed, endogenous expression levels. Unexpectedly, we found that codons decoded by rare transfer RNAs do not lead to slow translation under nutrient-rich conditions. Instead, Shine–Dalgarno-(SD)[6]-like features within coding sequences cause pervasive translational pausing. Using an orthogonal ribosome[7,8] possessing an altered anti-SD sequence, we show that pausing is due to hybridization between the mRNA and 16S ribosomal RNA of the translating ribosome. In protein-coding sequences, internal SD sequences are disfavoured, which leads to biased usage, avoiding codons and codon pairs that resemble canonical SD sites. Our results indicate that internal SD-like sequences are a major determinant of translation rates and a global driving force for the coding of bacterial genomes.**

Our current understanding of sequence-dependent translation rates *in vivo* derives largely from pioneering work begun in the 1980s[9–13]. These studies, which measured protein synthesis times using pulse labelling, established that different mRNAs could be translated with different elongation rates. In particular, messages decoded by less abundant tRNAs were translated slowly, although this effect was exaggerated by the overexpression of mRNA, which can lead to the depletion of available tRNAs[10]. Even with fixed tRNA usage, different synonymously coded mRNAs were translated at different rates[13]. This result, together with the observation of biased occurrence of adjacent codon pairs[14], suggested that tRNA abundance is not the only determinant of elongation rates. Further investigations addressing what determines the rate of translation *in vivo*, however, have been hampered by the limited temporal and positional resolution of existing techniques.

To provide a high-resolution view of local translation rates, we used the recently developed ribosome profiling strategy[3–5] to map ribosome occupancy along each mRNA (Supplementary Fig. 1). We focused on two distantly related bacterial species, the Gram-negative bacterium *Escherichia coli* and the Gram-positive bacterium *Bacillus subtilis*. To preserve the state of translation, cells were flash-frozen in liquid nitrogen after the rapid filtration of exponential-phase cultures. Ribosome-protected footprints were generated through nuclease treatment of cell extract in the presence of inhibitors of translation elongation (see Methods). These steps ensured that most ribosomes were polysome-associated after lysis and stayed assembled as 70S particles during digestion (Supplementary Fig. 2). After deep sequencing, 2,257 genes in *E. coli* and 1,580 genes in *B. subtilis* had an average coverage of at least ten sequencing reads per codon. The observed variability of ribosome footprint profiles across individual genes was highly reproducible ($r = 0.99$ between biological replicates; Supplementary Fig. 3).

Several observations argued that ribosome transit time is proportional to the occupancy at each position. First, we observed negligible internal initiation and early termination associated with ribosome pause sites (Supplementary Fig. 4). Second, ribosomes remained intact during footprinting, which enabled the large majority of ribosome-protected fragments to be captured (Supplementary Fig. 2). Third, the variability introduced during the conversion of RNA fragments into a sequenceable DNA library contributed minimally to our measures of variability in ribosome occupancy (Supplementary Fig. 5).

With our genome-wide view of local translation rates, we confirmed established examples of peptide-mediated stalling at transcripts *secM*[15] and *tnaC*[16] in *E. coli* and *mifM*[17] in *B. subtilis* (Fig. 1a and Supplementary Fig. 6). Strikingly, in addition to these known stalling sites, the observed ribosome occupancy was highly variable across coding regions, as illustrated for *secA* in Fig. 1a. We found that ribosome density often surpasses by more than tenfold the mean density, and the vast majority of these translational pauses are uncharacterized.

We first sought to determine whether the identity of the codon being decoded could account for the differences in local translation rates, by examining the average ribosome occupancy for each of the 61 codons in the ribosomal A-site. Surprisingly, there was little correlation between the average occupancy of a codon and existing measurements of the abundance of corresponding tRNAs[18] (Fig. 1b, c and Supplementary Fig. 7). Most notably, the six serine codons had the highest ribosome occupancy for *E. coli* cultured in Luria broth (Fig. 1b). Because serine is the first amino acid to be catabolized by *E. coli* when sugar is absent[19,20], we reasoned that the increased ribosome occupancy might be due to limited serine supply. Indeed, serine-associated pauses were greatly decreased in glucose-supplemented MOPS medium (Fig. 1c). The increase in serine codon occupancy when glucose becomes limiting confirmed our ability to capture translation rates at each codon. However, the identity of the A-site codon, which had less than a twofold effect on ribosome occupancy (Fig. 1c), could not account for the large variability in ribosome density along messages.

What, then, are the sequence features that cause slow translation? Without a priori knowledge about where such features would be located relative to the ribosomal A-site, we calculated the cross-correlation function between intragenic ribosome occupancy profiles and the presence of a given trinucleotide sequence on the mRNA independently of reading frames. Strong correlation was observed for six trinucleotide sequences (Fig. 1d) that resembled features found in Shine–Dalgarno (SD)[6] sequences. The highest correlation occurred when the SD-like feature was 8–11 nucleotides upstream of the position occupied by the ribosomal A-site. This spacing coincides with the optimal spacing for ribosome binding at start codons[21]. However, unlike canonical SD sites, which enable initiation of translation, the observed pauses were associated with SD-like features within the body of coding regions. The accumulation of ribosomes at internal SD-like sequences was observed across two divergent phyla of bacteria (Fig. 2a), suggesting that the phenomenon occurs generally in bacteria.

[1]Department of Cellular and Molecular Pharmacology, Howard Hughes Medical Institute, University of California, San Francisco, California 94158, USA.

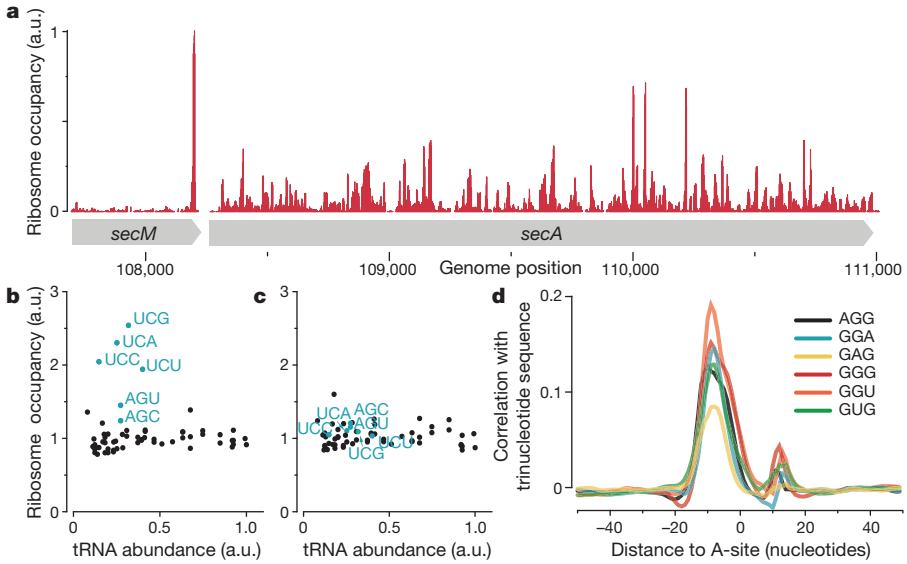


**Figure 1 | Analysis of translational pausing using ribosome profiling in bacteria.** **a**, Validation of the ribosome stalling site in the *secM* mRNA. **b**, **c**, Average ribosome occupancy of each codon relative to their respective tRNA abundances for *E. coli*. **b**, For growth in Luria broth, elevated occupancy at serine codons (blue) probably reflects preferential depletion of this amino acid. **c**, In glucose-rich medium, the ribosome occupancy is independent of tRNA abundance. **d**, Plot of cross-correlation function between ribosome occupancy profiles and the presence of the indicated trinucleotide sequences for *E. coli*.

The same correlation was not observed for the budding yeast *Saccharomyces cerevisiae*, whose ribosomes, like those of other eukaryotes, do not contain an anti-SD (aSD) site.

As predicted by a model in which the interaction between mRNA and the aSD site of the 16S rRNA drives pausing, the predicted hybridization free energy of a hexanucleotide to the aSD sequence was a strong indicator of its average downstream ribosome occupancy (Fig. 2b). Furthermore, there was a clear correspondence on individual transcripts between SD-like sequences and pauses. For example, Fig. 2c shows that in *ompF*, individual SD-like features are associated with elevated ribosome occupancy 8–11 nucleotides downstream. Moreover, a synonymous mutation (GGUGGU to GGCGGC) that decreased the affinity for the aSD site led to reduced ribosome occupancy specifically at the mutated sequence, suggesting a causal relationship between the SD-like feature and the excess ribosome density.

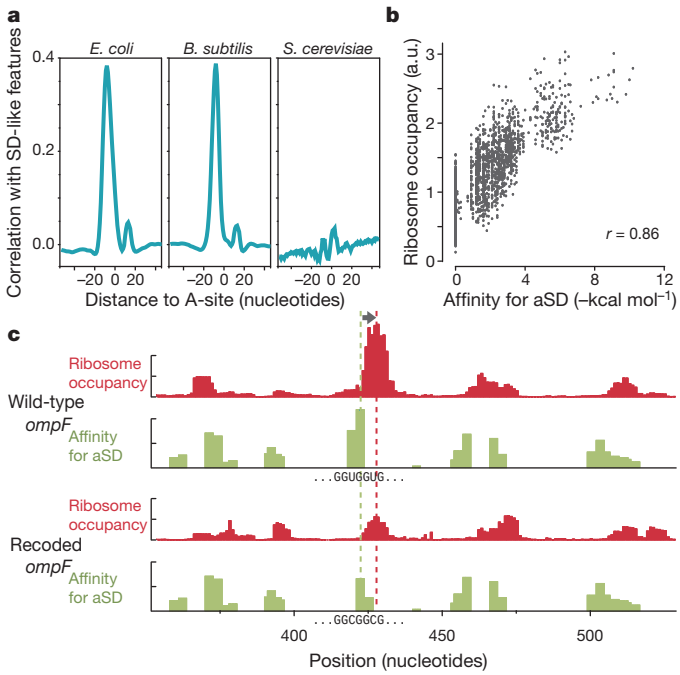


**Figure 2 | Relationship between ribosome pausing and internal Shine–Dalgarno sequences.** **a**, Plot of correlation between ribosome occupancy and SD-like features for *E. coli, B. subtilis* and *S. cerevisiae*. **b**, Plot of the average ribosome occupancy of hexanucleotide sequences relative to their affinity for the anti-Shine–Dalgarno sequence. **c**, Reprogrammed pausing by recoding the *ompF* mRNA. Ribosome occupancy (red) increases when the A-site is 8–11 nucleotides downstream (arrow) of SD-like features (green). Synonymous mutations replacing the SD-like sequence (GGUGGUG) in wild-type *ompF* (top) with a sequence (GGCGGCG) with lower affinity for the aSD site (bottom) caused a corresponding decrease in ribosome pausing.

We next sought to evaluate directly whether the excess footprint density seen at internal SD-like sequences was due to pausing of elongating ribosomes rather than attempted internal initiation, driven by SD–aSD interactions (Fig. 3a). To distinguish between these possibilities, we used a previously described orthogonal ribosome (O-ribosome) system in which a mutant form of the 16S rRNA with an altered aSD site is expressed together with wild-type 16S rRNA[8]. O-ribosomes containing the mutant 16S RNA will only translate a target mRNA that has the corresponding orthogonal SD (O-SD) sequence before its start codon. Conversely, a message whose translation is driven by the O-SD sequence will only be translated by O-ribosomes, and not by wild-type ribosomes. This system thus allows one to determine the source of regions of excess ribosome footprints, because elongating O-ribosomes would pause at internal O-SD sequences, whereas attempted internal initiation would still occur at SD sequences as a result of the cellular pool of wild-type ribosomes.

We compared the ribosome occupancy profiles of a *lacZ* message that was translated by either O-ribosomes or wild-type ribosomes. The occupancy profile of the *lacZ* message exclusively translated by O-ribosomes was correlated with O-SD-like features, and not with SD-like features (Fig. 3c). This is in marked contrast with the same *lacZ* sequence translated by wild-type ribosomes (Fig. 3b). As an internal control in O-ribosome-expressing cells, all other genes, which were translated by wild-type ribosomes, still maintained SD-correlated ribosome occupancy profiles (Fig. 3c). These observations established that elongating ribosomes pause during protein synthesis and that hybridization between the aSD site in the elongating ribosome and internal SD-like sequences gives rise to these pauses.

Global analysis of pause sites revealed that internal SD-like sequences are the dominant feature controlling translational pausing: about 70% of the strong pauses (that is, those that have ribosome

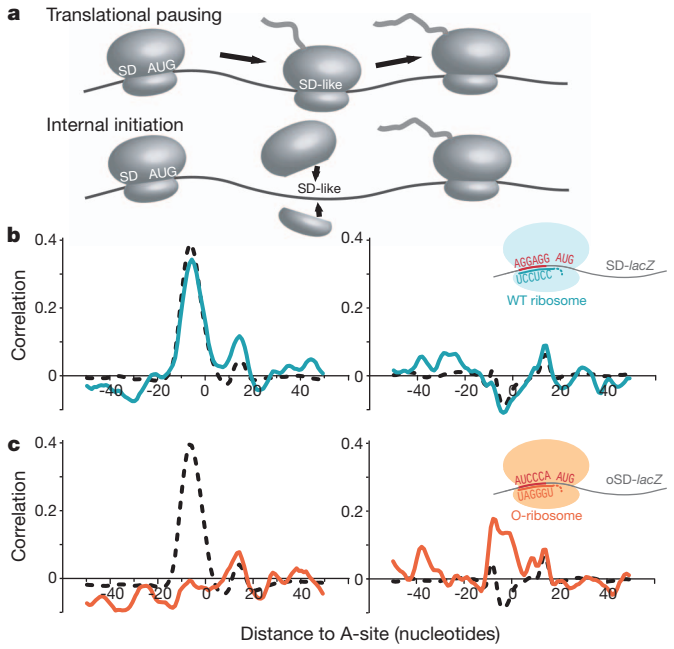


**Figure 3 | Pausing of elongating ribosomes due to SD–aSD interaction.** **a**, Two models could account for the excess ribosome density at internal SD-like sequences. **b**, Ribosome occupancy of *lacZ* mRNA translated by wild-type ribosome. Like other genes translated by the wild-type ribosome, the ribosome occupancy pattern along *lacZ* is correlated with the presence of SD-like sequences (left), not with the O-SD sequence (right). Cyan, *lacZ*; black; all other genes. **c**, Ribosome occupancy of *lacZ* mRNA translated by orthogonal ribosome (O-ribosome). Unlike other genes in the same cells, the specialized O-SD *lacZ* has ribosome pausing at internal O-SD-like sequences (right), not at SD-like sites (left). Orange, *lacZ*; black; all other genes.

occupancies more than tenfold over the mean) are associated with SD sites (Supplementary Fig. 8). Although the interaction between internal SD sequences in a message and elongating ribosomes has been documented in specialized cases, including promoting frame-shifting *in vivo*[22,23] and ribosome stalling in single-molecule experiments *in vitro*[24], there was little indication that internal SD-like sequences are the major determinant of elongation rate during translation.

Because translational pausing limits the amount of free ribosomes available for initiating protein synthesis, widespread internal SD-like sequences could decrease bacterial growth rates. Accordingly, we found that strong SD-like sequences are generally avoided in the coding region of *E. coli* genes: hexamer sequences that strongly bind aSD sites are universally rare, whereas low-affinity hexamers have variable rates of occurrence (Fig. 4a). Consistent with translational pausing being the driving force for this bias, depletion of SD-like sequences was observed only in protein-coding genes, and not in genes encoding rRNA or tRNA (Supplementary Fig. 9). The selection against SD-like features in turn impacts both synonymous codon choice and codon-pair choice. At the codon level, SD-like codons GAG, AGG and GGG are all minor codons in *E. coli* and *B. subtilis*. The evolutionary origin of codon selection is often attributed to differences in tRNA abundance[2,25] because its level is correlated with codon usage[18]. Instead, we propose that SD-like codons are disfavoured as a result of their interactions with rRNA, and that tRNA expression levels followed codon adaptation.

At the codon-pair level, we can now account for the selection against two consecutive codons that resemble SD sequences. This is illustrated for Gly-Gly pairs, which are coded by GGNGGN sequences (Fig. 4b). The most abundant Gly-Gly coding sequence, GGCGGC, has the lowest affinity for the aSD sequence, whereas Gly-Gly coding sequences that strongly resemble SD sites, including GGAGGU, which perfectly complements the aSD site, rarely appear. This under-representation holds even after correcting for the usage of individual codons (Fig. 4b, colour coding); for example, GGAGGU is considerably less common than GGUGGA. Other amino-acid pairs that can be coded with strong SD sites also showed the same bias (Supplementary Fig. 10). The preference in codon pairs stems from the sequence identity and not codon identity, because the same trend is seen in hexamers that are not aligned to codon pairs (Supplementary Fig. 11). Although not every bias in codon-pair usage can be explained here, the disadvantage associated with SD-induced translational pausing offers a clear mechanistic view of why certain codon pairs are avoided.

Despite the selection against internal SD-like sequences, they remain a major driving force for translational pausing. In addition, we found similar pausing patterns between conserved genes in *E. coli* and *B. subtilis* (Fig. 4c). For an mRNA encoding a specific protein, it may not be possible to fully eliminate sequences with affinity for the aSD site without changing the peptide sequence. For example, in the case of Gly-Gly, even the GGCGGC pair has substantial affinity for the aSD site. The optimization for translation rate therefore cannot be achieved only at the level of mRNA coding: it is also constrained by the requirement to make a functional peptide sequence.

The observation that the ability of elongating ribosomes to interact with SD-like sequences is highly conserved suggests that this mechanism of pausing is exploited for functional purposes. Indeed, a highly conserved internal SD site exists in the gene encoding peptide chain release factor 2 (RF2)[26]. This sequence has an important function in promoting a translational frameshift to enable its expression. In addition, pausing at internal SD-like sites could modulate the co-translational folding of the nascent peptide chain (Supplementary Fig. 12). Finally, given the coupling between transcription and translation in bacteria[27,28], pausing

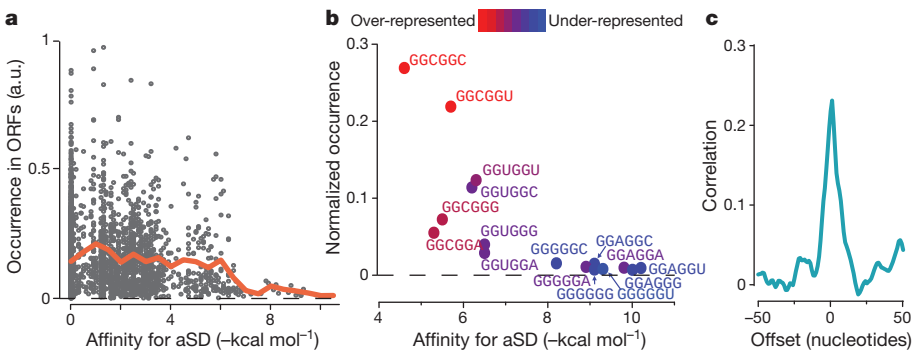


**Figure 4 | Selection against SD-like sequences and the constraint on protein coding.** **a**, Rate of occurrence of hexanucleotide sequences in *E. coli* messages relative to their predicted affinity for the aSD site. The orange line shows the average occurrence within a bin size of 0.5 kcal mol$^{-1}$. **b**, Occurrence of codon-pairs for Gly-Gly residues relative to their predicted affinity for the aSD site. The colour coding represents the enrichment in occurrence of codon pairs after correcting for the usage of single codons. **c**, Cross-correlation function of ribosome occupancy profiles between conserved genes in *E. coli* and *B. subtilis*. Zero offset means that the two sequences are aligned at each amino-acid residue.

at SD sites could be exploited for transcriptional regulation. We observed internal SD sites and pausing near the stop codon of transcription attenuation leader peptides[29], including *trpL* and *thrL* (Supplementary Fig. 13). In contrast to ribosome stalling at regulatory codons during starvation, slow translation near the stop codon could protect alternative structural mRNA elements to prevent the formation of anti-termination stem-loops, thereby ensuring proper transcription termination[30]. Our approach and the genome-wide data lay the groundwork for further gene-specific functional studies of translational pausing.

From a more practical perspective, ribosome pausing at internal SD sites presents both a challenge and an opportunity for heterologous protein expression in bacteria. Overexpression of eukaryotic proteins with strong internal SD sites would sequester ribosomes and compromise protein yield. Internal SD sequences could be reduced by recoding the gene, which has not been considered in conventional strategies of simple codon optimization or overexpression of rare tRNAs. Conversely, recoding can introduce internal SD sites if pausing is required for co-translational processing. Positioning of internal SD sites therefore adds another dimension to the optimization of heterologous protein expression.

## METHODS SUMMARY

*E. coli* MG1655 and *B. subtilis* 168 were used as wild-type strains. *E. coli* BJW9 has synonymous substitutions at G141 and G142 in the *ompF* gene. The orthogonal ribosome experiment was performed in *E. coli* BW25113 with two plasmids: pSC101-G9, expressing orthogonal 16S rRNA, and pJW1422, expressing O-SD-*lacZ* mRNA. pSC101-G9 was a gift from J. Chin[8]. pJW1422 has *lacZ* driven from a *tacII* promoter and an O-ribosome binding site 5′-AUCCCA-3′. Luria broth was used for *B. subtilis* culture. Cell cultures were harvested at a $D_{600}$ of 0.3–0.4. Flash-freezing and ribosome footprinting was described previously[5]. 5′-Guanylyl imidodiphosphate (3 mM) was added to the lysate before thawing and during footprinting to prevent translation after lysis. Conversion of mRNA footprints to a complementary DNA library was described previously[4,5]. Deep sequencing was performed on an Illumina HiSeq 2000 system, and the results were aligned to reference genomes using Bowtie v. 0.12.0. The cross-correlation function is defined as

$$C_i = \frac{\langle x_{j+i} y_j \rangle - \mu_X \mu_Y}{\sigma_X \sigma_Y}$$

for the series $X = x_1, x_2, \ldots, x_N$ and $Y = y_1, y_2, \ldots, y_N$, where $\mu_X$ and $\sigma_X$ are the average and the standard deviation of series $X$, respectively.

**Full Methods** and any associated references are available in the online version of the paper at www.nature.com/nature.

**Received 21 November 2011; accepted 16 February 2012.**
**Published online 28 March 2012.**

1. Kramer, G., Boehringer, D., Ban, N. & Bukau, B. The ribosome as a platform for co-translational processing, folding and targeting of newly synthesized proteins. *Nature Struct. Mol. Biol.* **16,** 589–597 (2009).
2. Plotkin, J. B. & Kudla, G. Synonymous but not the same: the causes and consequences of codon bias. *Nature Rev. Genet.* **12,** 32–42 (2011).
3. Ingolia, N. T., Ghaemmaghami, S., Newman, J. R. & Weissman, J. S. Genome-wide analysis *in vivo* of translation with nucleotide resolution using ribosome profiling. *Science* **324,** 218–223 (2009).
4. Ingolia, N. T., Lareau, L. F. & Weissman, J. S. Ribosome profiling of mouse embryonic stem cells reveals the complexity and dynamics of mammalian proteomes. *Cell* **147,** 789–802 (2011).
5. Oh, E. *et al.* Selective ribosome profiling reveals the cotranslational chaperone action of trigger factor *in vivo*. *Cell* **147,** 1295–1308 (2011).
6. Shine, J. & Dalgarno, L. The 3′-terminal sequence of *Escherichia coli* 16S ribosomal RNA: complementarity to nonsense triplets and ribosome binding sites. *Proc. Natl Acad. Sci. USA* **71,** 1342–1346 (1974).
7. Hui, A. & de Boer, H. A. Specialized ribosome system: preferential translation of a single mRNA species by a subpopulation of mutated ribosomes in *Escherichia coli*. *Proc. Natl Acad. Sci. USA* **84,** 4762–4766 (1987).
8. Rackham, O. & Chin, J. W. A network of orthogonal ribosome•mRNA pairs. *Nature Chem. Biol.* **1,** 159–166 (2005).
9. Varenne, S., Buc, J., Lloubes, R. & Lazdunski, C. Translation is a non-uniform process. Effect of tRNA availability on the rate of elongation of nascent polypeptide chains. *J. Mol. Biol.* **180,** 549–576 (1984).
10. Pedersen, S. *Escherichia coli* ribosomes translate *in vivo* with variable rate. *EMBO J.* **3,** 2895–2898 (1984).
11. Sorensen, M. A., Kurland, C. G. & Pedersen, S. Codon usage determines translation rate in *Escherichia coli*. *J. Mol. Biol.* **207,** 365–377 (1989).
12. Andersson, S. G. & Kurland, C. G. Codon preferences in free-living microorganisms. *Microbiol. Rev.* **54,** 198–210 (1990).
13. Sorensen, M. A. & Pedersen, S. Absolute *in vivo* translation rates of individual codons in *Escherichia coli*. The two glutamic acid codons GAA and GAG are translated with a threefold difference in rate. *J. Mol. Biol.* **222,** 265–280 (1991).
14. Gutman, G. A. & Hatfield, G. W. Nonrandom utilization of codon pairs in *Escherichia coli*. *Proc. Natl Acad. Sci. USA* **86,** 3699–3703 (1989).
15. Nakatogawa, H. & Ito, K. The ribosomal exit tunnel functions as a discriminating gate. *Cell* **108,** 629–636 (2002).
16. Gong, F. & Yanofsky, C. Instruction of translating ribosome by nascent peptide. *Science* **297,** 1864–1867 (2002).
17. Chiba, S. *et al.* Recruitment of a species-specific translational arrest module to monitor different cellular processes. *Proc. Natl Acad. Sci. USA* **108,** 6073–6078 (2011).
18. Dong, H., Nilsson, L. & Kurland, C. G. Co-variation of tRNA abundance and codon usage in *Escherichia coli* at different growth rates. *J. Mol. Biol.* **260,** 649–663 (1996).
19. Pruss, B. M., Nelms, J. M., Park, C. & Wolfe, A. J. Mutations in NADH:ubiquinone oxidoreductase of *Escherichia coli* affect growth on mixed amino acids. *J. Bacteriol.* **176,** 2143–2150 (1994).
20. Sezonov, G., Joseleau-Petit, D. & D'Ari, R. *Escherichia coli* physiology in Luria–Bertani broth. *J. Bacteriol.* **189,** 8746–8749 (2007).
21. Chen, H., Bjerknes, M., Kumar, R. & Jay, E. Determination of the optimal aligned spacing between the Shine–Dalgarno sequence and the translation initiation codon of *Escherichia coli* mRNAs. *Nucleic Acids Res.* **22,** 4953–4957 (1994).
22. Weiss, R. B., Dunn, D. M., Dahlberg, A. E., Atkins, J. F. & Gesteland, R. F. Reading frame switch caused by base-pair formation between the 3′ end of 16S rRNA and the mRNA during elongation of protein synthesis in *Escherichia coli*. *EMBO J.* **7,** 1503–1507 (1988).
23. Larsen, B., Wills, N. M., Gesteland, R. F. & Atkins, J. F. rRNA–mRNA base pairing stimulates a programmed −1 ribosomal frameshift. *J. Bacteriol.* **176,** 6842–6851 (1994).
24. Wen, J. D. *et al.* Following translation by single ribosomes one codon at a time. *Nature* **452,** 598–603 (2008).
25. Ikemura, T. Correlation between the abundance of *Escherichia coli* transfer RNAs and the occurrence of the respective codons in its protein genes: a proposal for a synonymous codon choice that is optimal for the *E.coli* translational system. *J. Mol. Biol.* **151,** 389–409 (1981).
26. Baranov, P. V., Gesteland, R. F. & Atkins, J. F. Release factor 2 frameshifting sites in different bacteria. *EMBO Rep.* **3,** 373–377 (2002).
27. Burmann, B. M. *et al.* A NusE:NusG complex links transcription and translation. *Science* **328,** 501–504 (2010).
28. Proshkin, S., Rahmouni, A. R., Mironov, A. & Nudler, E. Cooperation between translating ribosomes and RNA polymerase in transcription elongation. *Science* **328,** 504–508 (2010).
29. Kolter, R. & Yanofsky, C. Attenuation in amino acid biosynthetic operons. *Annu. Rev. Genet.* **16,** 113–134 (1982).
30. Elf, J. & Ehrenberg, M. What makes ribosome-mediated transcriptional attenuation sensitive to amino acid limitation? *PLOS Comput. Biol.* **1,** e2 (2005).

**Supplementary Information** is linked to the online version of the paper at www.nature.com/nature.

**Acknowledgements** We thank E. Reuman, D. Burkhardt, C. Jan, C. Gross, J. Elf and members of the Weissman laboratory for discussions; J. Dunn for ribosome profiling data on *S. cerevisiae*; C. Chu for help with sequencing; and J. Chin for orthogonal ribosome reagents and advice. This research was supported by the Helen Hay Whitney Foundation (to G.W.L.) and by the Howard Hughes Medical Institute (to J.S.W.).

**Author Contributions** G.W.L. and J.S.W. designed the experiments. G.W.L. performed experiments and analysed the data. E.O. provided technical support and preliminary data. G.W.L. and J.S.W. wrote the manuscript.

**Author Information** The footprint sequencing data are deposited in the Gene Expression Omnibus (GEO) under accession number GSE35641. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to J.S.W. (weissman@cmp.ucsf.edu).