Cell Reports
Report

Differential Translation Tunes Uneven Production of Operon-Encoded Proteins

Tessa E.F. Quax,[1,*] Yuri I. Wolf,[2] Jasper J. Koehorst,[1] Omri Wurtzel,[3] Richard van der Oost,[1] Wenqi Ran,[2] Fabian Blombach,[1] Kira S. Makarova,[2] Stan J.J. Brouns,[1] Anthony C. Forster,[4] E. Gerhart H. Wagner,[4] Rotem Sorek,[3] Eugene V. Koonin,[2] and John van der Oost[1,*]

[1]Laboratory of Microbiology, Wageningen University, Dreijenplein 10, 6703 HB Wageningen, The Netherlands
[2]National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda, MD 20894, USA
[3]Department of Molecular Genetics, Weizmann Institute of Science, Rehovot 76100, Israel
[4]Department of Cell and Molecular Biology, Uppsala University, Husargatan 3, Box 596, Uppsala 75124, Sweden
*Correspondence: tessa.quax@wur.nl (T.E.F.Q.), john.vanderoost@wur.nl (J.v.d.O.)
http://dx.doi.org/10.1016/j.celrep.2013.07.049

SUMMARY

Clustering of functionally related genes in operons allows for coregulated gene expression in prokaryotes. This is advantageous when equal amounts of gene products are required. Production of protein complexes with an uneven stoichiometry, however, requires tuning mechanisms to generate subunits in appropriate relative quantities. Using comparative genomic analysis, we show that differential translation is a key determinant of modulated expression of genes clustered in operons and that codon bias generally is the best in silico indicator of unequal protein production. Variable ribosome density profiles of polycistronic transcripts correlate strongly with differential translation patterns. In addition, we provide experimental evidence that de novo initiation of translation can occur at intercistronic sites, allowing for differential translation of any gene irrespective of its position on a polycistronic messenger. Thus, modulation of translation efficiency appears to be a universal mode of control in bacteria and archaea that allows for differential production of operon-encoded proteins.

INTRODUCTION

The operon concept was developed over 50 years ago by Jacob and Monod (1961). Their pioneering analyses revealed a fundamental and characteristic feature of prokaryotic genome organization, i.e., clustering of functionally related genes (Koonin, 2009). The operon organization allows for coregulated gene expression (Brenner et al., 1961; French et al., 2007; Grundy and Henkin, 2006). This is evidently advantageous when equimolar amounts of gene products are required, for instance, to generate multisubunit complexes with an even stoichiometry. However, a substantial number of operon-encoded multisubunit complexes have an uneven stoichiometry and many of these complexes play key roles in cellular processes such as protein translation, secretion, energy conservation, and antiviral defense (Abrahams et al., 1994; Dunkle et al., 2011; Johnson et al., 2006; Jore et al., 2011). Although it is anticipated that a specific tuning mechanism is required to generate subunits of these complexes in appropriate relative quantities, the elucidation of its molecular basis is a long-standing issue.

Control of subunit stoichiometry theoretically can be established at three levels: transcription, translation, and/or protein turnover. Only a few proteolysis substrates have been recognized to date, and comparison of the protein degradation rates awaits the generation of comprehensive proteomic pulse-chase databases (Gur et al., 2011). Although a contribution of different protein half-life values cannot be ruled out, it is considered most likely that prokaryotes avoid substantial energy loss by controlling different rates of subunit biosynthesis in order to obtain the appropriate relative quantities. Hence, uneven subunit stoichiometry of operon-encoded protein complexes is likely to be controlled by fine-tuning of differential transcription and/or translation rates. In the classical operon model, multiple genes/cistrons are transcribed on a single polycistronic messenger, resulting in the same levels of messenger RNA (mRNA) segments that are part of the operon mRNA (Jacob and Monod, 1961). Indeed, many full-length polycistronic mRNAs have been identified experimentally, and in these cases differential transcription or stability cannot account for uneven protein output. The recent development of whole-transcriptome sequencing has allowed for a more detailed analysis of operon transcription, which has revealed widespread internal transcription initiation and/or termination sites (Koide et al., 2009; Wurtzel et al., 2010).

Another possible means of achieving differential production of operon-encoded proteins involves regulation of translation efficiency. Translation efficiency depends on a range of features hidden in the noncoding and coding fragments of the transcripts' nucleotide sequences, and tuning may occur at the level of translation initiation and translation elongation (Cannarozzi et al., 2010; Drummond and Wilke, 2008; Fredrick and Ibba, 2010; Gingold and Pilpel, 2011; Kudla et al., 2009; Li et al., 2012; Shao et al., 2012; Sharp and Li, 1987; Stenström et al., 2001; Timmermans and Van Melderen, 2010; Tuller et al., 2010).

Aside from analyzing the relative production of the polypeptide end products, until recently, no high-throughput methods have been available for the direct monitoring of translation efficiency. The recent development of a ribosome density profiling method has provided insight into transcriptome-wide translation efficiency and offers the exciting possibility to study operon-encoded protein translation in greater detail (Ingolia et al., 2009; Li et al., 2012). In this report, we present results of a comparative genomic analysis of operon-encoded proteins showing that differential translation is a key determinant of the modulated expression of individual genes that are part of operons.

RESULTS

Selection of Data Sets

Expression of a given gene is influenced by many determinants, each of which plays a role in producing appropriate levels of the encoded proteins (Goldberger et al., 1976; Kaberdin and Bläsi, 2006; Marzi et al., 2007; Timmermans and Van Melderen, 2010). To understand how differential production of operon-encoded proteins is achieved, one should assess the relative contribution of each of these key controlling features to the expression of the individual genes within an operon. In addition to the rapidly growing database of archaeal and bacterial genomes (http://www.ncbi.nlm.nih.gov/genome/), recently compiled prokaryotic transcriptome data sets (Cho et al., 2009; Mendoza-Vargas et al., 2009; Toledo-Arana et al., 2009; Wurtzel et al., 2010, 2012a, 2012b) and bacterial ribosome density profiles (Ingolia et al., 2009; Li et al., 2012) are providing the necessary data to address the issue of tuning uneven subunit stoichiometry.

A set of well-conserved operon-encoded protein complexes from prokaryotes was selected to allow for the identification of factors that correlate best with, and thus may be causative for, differential protein production. Ten operons were chosen on the basis of their established uneven protein stoichiometry, and because they are conserved in many bacterial or archaeal genomes. In addition, two operons were included as controls because they encode complexes in which all subunits are present in equal amounts (Tables S1 and S2). We performed comparative analyses on a set of 1,055 bacterial and archaeal genomes, from which we selected a subset of 383 to avoid biases caused by the close relationships among some of the available genomes, e.g., at the species or subspecies level (Tables S1 and S2). All of the selected operons encode protein complexes that play roles in important cellular processes (e.g., translation, secretion, and energy production (Abrahams et al., 1994; Beyenbach and Wieczorek, 2006; Dunkle et al., 2011; Efremov et al., 2010; Errington, 2003; Ghosh and Albers, 2011; Jakob et al., 2009; Johnson et al., 2006; Jore et al., 2011; Wiedenheft et al., 2011).

Differential Transcription

We first assessed whether differential transcription or, more specifically, different levels of mRNA segments encoding the cistrons in the selected set of operons reflect the differential production of subunits of complexes with uneven stoichiometry. To this end, we used high-throughput whole-transcriptome sequencing data and tiling-array expression data for representative microbes, including three bacteria and an archaeon (Cho et al., 2009; Mendoza-Vargas et al., 2009; Toledo-Arana et al., 2009; Wurtzel et al., 2010, 2012a, 2012b). Differences in mRNA levels corresponding to genes within an operon could arise from internal transcription initiation/termination and/or processing, and differential decay of polycistronic mRNA (Li and Altman, 2004). In addition, we analyzed experimentally determined transcription start site (TSS) maps for each organism to identify potential alternative transcriptional units within the selected clusters. In the majority of the analyzed operons (90%), deep transcriptome sequencing data showed that genes encoding different proteins were transcribed to similar levels (Figure S1; Table S1). Thus, modulation at the transcription level, i.e., generating nonstoichiometric mRNA segment levels, may contribute to some extent, but does not appear to be a dominant factor in tuning the differential production of proteins.

Differential Translation—In Silico Analysis

As the minor effect observed at the transcription level cannot account for drastically different protein output from operons, we set out to analyze the contribution of differential translation. Previous analyses of (monocistronic) transcripts in prokaryotes have shown that several factors could contribute to the overall efficiency of the translation process (Kudla et al., 2009; Cannarozzi et al., 2010; Tuller et al., 2010; Sharp and Li, 1987). We analyzed the correlation of each of the factors with protein subunit stoichiometry, using the aforementioned data set of well-characterized complexes.

Translation Initiation

Conventional translation initiation in bacteria involves binding of the 30S ribosomal subunit to the ribosome-binding site (RBS) of an mRNA. This is generally dependent on the Shine-Dalgarno (SD) sequence, which base pairs with the anti-SD sequence in 16S ribosomal RNA (rRNA) to guide selection of the correct start codon. The rate of translation initiation depends on (1) the strength of the interaction of SD/anti-SD base pairing (Vellanoweth and Rabinowitz, 1992) and (2) the accessibility of the RBS (involving primarily SD and/or the start codon), which is negatively affected by stable secondary structure (Kudla et al., 2009). In the absence of a canonical SD motif, the codon following the initiation codon may affect the translation initiation efficiency (Stenström et al., 2001).

When we calculated the RNA hybridization energy between the SD sequences and the anti-SD sequences (G) of the genes within the selected operons, the majority of the RBSs failed to reveal statistically significant differences between genes in operons (Table S2). Similarly, analysis of adenine enrichment of the second codon showed no association with the stoichiometry of the complex subunits (Table S2). These observations suggest that the affinity of the SD/anti-SD interaction and the nature of the second codon do not play major roles in differentiating translation efficiency.

Next, the propensity to form secondary RNA structure (E) in RBS regions of genes (−20 to +20 bp relative to the start

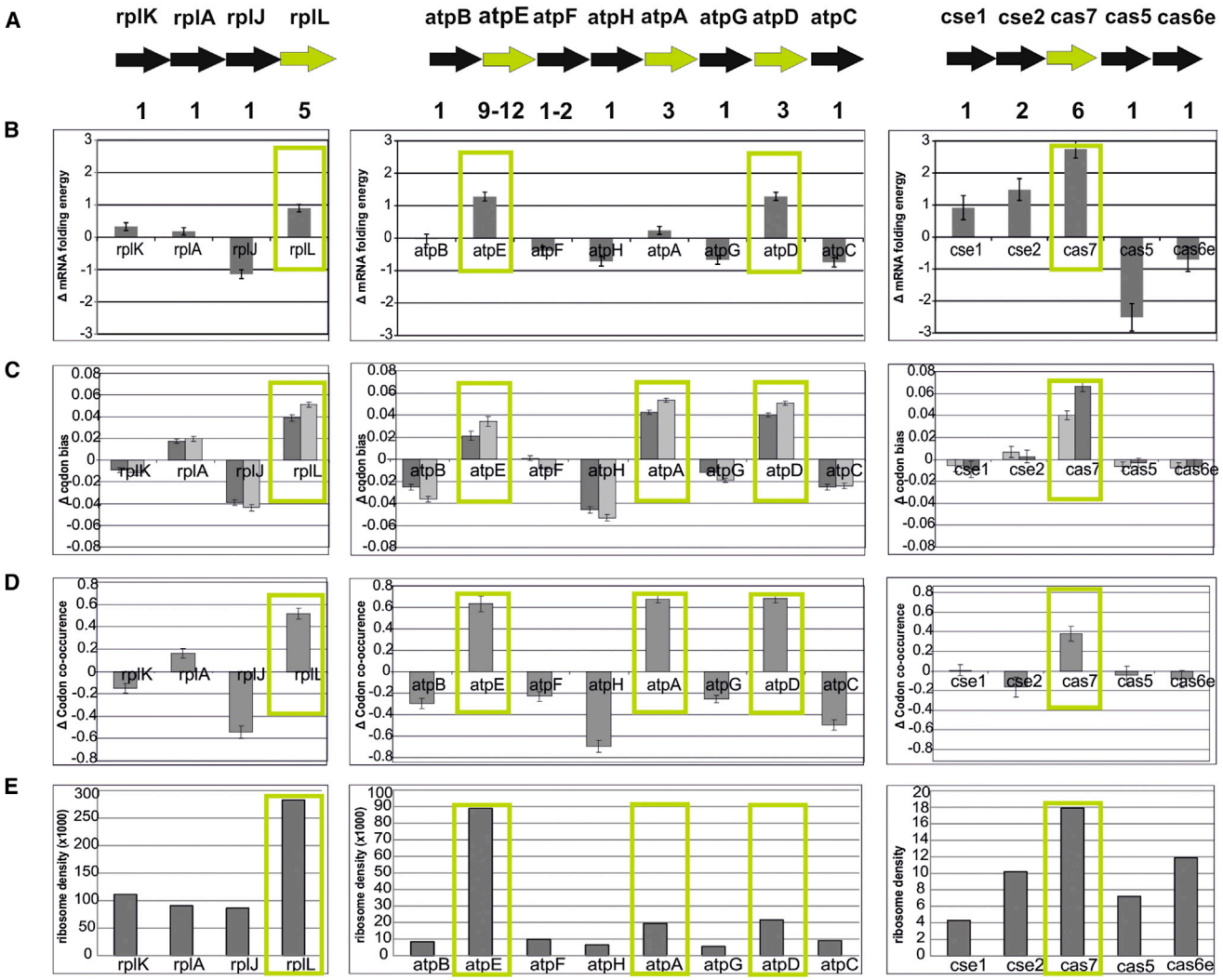

Figure 1. Cistron Properties Correlated with Stoichiometry of Operon-Encoded Protein Complexes

Analysis of the ribosomal protein operon L7/L12 (high expression), the F-type ATPase (moderate expression), and the type I-E Cascade complex (low expression).

(A) Selected operons (block arrows) and the stoichiometry in the corresponding protein complexes.

(B) Predicted mRNA folding energy (ΔE) of the RBS region of each cistron (−20 to +20 bp relative to the start codon).

(C) Codon bias; ΔF (optimal codon usage) is shown by dark gray and ΔCAI is shown by light gray.

(D) Codon co-occurrence (ΔCo).

(E) Ribosome density profiles per gene. The green arrows represent genes in each operon that encode the most abundant subunit(s), and green rectangles denote the corresponding positive deviations in codon bias (ΔF > 0.02), codon co-occurrence (highest value), low RNA folding potential (highest energy value), and/or ribosome density (highest value).

Error bars represent one unit of standard deviation. See also Figures S1, S2, S3, S4, S5, and Table S2.

codon) of selected operons was calculated. To compute the operon-specific values, accounting simultaneously for genome-specific and operon-specific biases, the mean for the respective operon was subtracted from individual gene values (ΔE value) and the differential was adduced over the complete set of operons of the given type. This analysis revealed a moderate but significant correlation between subunit stoichiometry and mRNA folding (ΔE: $r_{Spearman}$ = 0.57, p = 0.0092; Figures 1A, 1B, S2, and S3; Table S2). These results indicate that the structural accessibility of RBS might contribute to the differential gene expression in several of the selected operons.

Translation Elongation

The efficiency (accuracy and/or rate) of translation elongation depends on the coding sequence of a gene (Drummond and Wilke, 2008; Gingold and Pilpel, 2011). Codon bias reflects differences between isoacceptor transfer RNAs (tRNAs) with respect to abundance, amino acid charging, and kinetics (Forster, 2012; Fredrick and Ibba, 2010; Welch et al., 2009). In addition, translation is enhanced by the co-occurrence of isoaccepting codons (Cannarozzi et al., 2010; Shao et al., 2012), i.e., codons that are recognized by the same tRNA (by Watson-Crick base pairing or by wobbling, which allows non-Watson-Crick

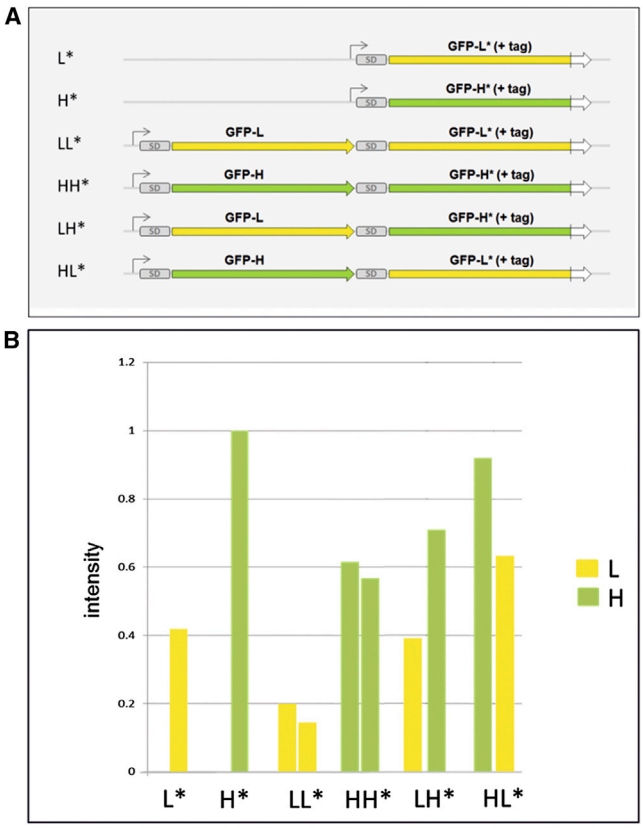

Figure 2. Translation Efficiency Influences Protein Expression from Individual Genes within Operons

(A) Expression constructs under control of the T7 promoter (gray arrow) encoding different combinations of two identical GFP polypeptides, with synonymous mutations, resulting in low (yellow block arrow, *gfp*L) or high (green block arrow, *gfp*H) translation efficiency. The single genes, as well as the downstream gene of the operons, are translational fusions to a Strep-tag (white block arrow; in construct name indicated by an asterisk). SD, Shine-Dalgarno sequence (Figure S6).

(B) Quantification of western blot with GFP antibody on whole-cell lysates of the variant GFP-expressing constructs after expression in *E. coli*. Equal total amounts of cellular proteins were loaded in order to allow comparison between different samples (Figure S6).

base pairing between two nucleotides of interacting RNA molecules). A comparative analysis of codon usage and co-occurrence was performed in genes of the selected operons (Table S2). For each analyzed genome, the classical Codon Adaptation Index (CAI) (Sharp and Li, 1987), the percentage of optimal codons (F) (Ikemura, 1981), and codon co-occurrence (Co) (Cannarozzi et al., 2010) were calculated. A significant correlation was observed between protein subunit stoichiometry and family-averaged scores for codon co-occurrence (ΔCo: $r_{Spearman}$ = 0.63, p = 0.0099) (Figures 1A, 1D, S2, and S3; Table S2). The results obtained for CAI and F were in excellent agreement and indicated that in all of the selected operons, the genes encoding the most abundant proteins (>2 copies per complex) contained the highest percentage of optimal codons and displayed the highest CAI values (Figures 1A, 1C, S2, and S3; Table S2). A threshold value of 0.02 for the family average value ΔF was determined to minimize the error rate for the prediction of high-stoichiometry subunits, resulting in a prediction accuracy of 96% for the analyzed set of operons (Figure S4). Overall, for all of the selected operons, there was a strong correlation between codon usage and protein stoichiometry (for ΔF: $r_{Spearman}$ = 0.75, p = 0.0002; for ΔCAI: $r_{Spearman}$ = 0.71, p = 0.0002; Experimental Procedures).

Ribosome Profiling

To determine the in vivo translation rates, we analyzed the ribosome density profiles of *Escherichia coli* and *Bacillus subtilis* that were recently reported by Li et al. (2012). The ribosome profiling strategy allows for quantitative monitoring of protein production in vivo because ribosome density values closely correlate with translation efficiencies if the crosslinked ribosomes are evenly distributed over the coding sequence (Ingolia et al., 2009). On the other hand, local peaks in density characterize sequences with lower translation efficiency due to ribosome stalling (Li et al., 2012). For the selected operons, we found a strong correlation of ribosome density with protein subunit stoichiometry (Figures 1A, 1E, S5; Table S2). This finding is in perfect agreement with our conclusion that uneven production of subunits of operon-encoded protein complexes is tuned by differential translation.

De Novo Translation Initiation

The fact that genes coding for abundant subunits do not necessarily occupy the first position in the respective operons (Figures 1, S2, S3, and S5) implies that internal translation initiation is required to allow for elevated translation rates of a downstream cistron. To investigate the possibility of internal initiation of translation, we used operons consisting of two synthetic GFP-encoding genes with synonymous mutations (*gfp*L and *gfp*H, with low and high expression, respectively, as reported by Kudla et al. [2009]; Figures 2A and S6). After induction of expression in *E. coli*, detection by western blot analysis showed that moderate levels of GFP variants were produced (Figures 2B and S6). When *gfp*L and *gfp*H were combined in one operon, the expression level of *gfp*H was substantially higher than that of *gfp*L, irrespective of its position in the operon (Figures 2B and S6), confirming that the position of a cistron in the operon does not appreciably affect protein expression levels. We conclude that the differential protein production of GFP-L and GFP-H resulted from differences in translation efficiency, in accord with internal de novo translation initiation.

DISCUSSION

In line with the definition of an operon, our analyses indicate that the differential expression of individual genes in the selected operons shows only a limited dependence on differential transcription. Rather, we provide evidence that expression of genes in operons is predominantly controlled at the translation level. Although protein degradation may not be the most economical method for balancing protein ratios, we cannot rule out its importance.

Elevated overall translation of a cistron, potentially at any position on a polycistronic mRNA, requires enhanced translation

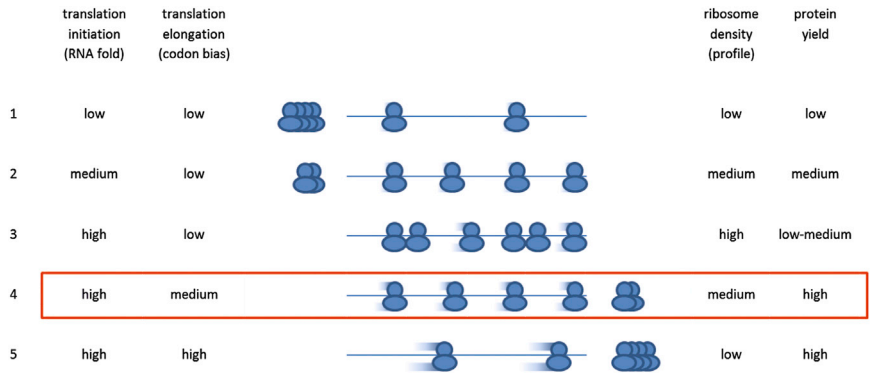

Figure 3. Models Describing Tuning of the Translation Process

Five scenarios are depicted with different rates of translation initiation and/or translation elongation. The corresponding ribosome density profiles and the expected protein yield are based on the assumption that differential transcription is insignificant (as has been demonstrated experimentally for the majority of the operons analyzed in this study). Scenario 1: relatively little protein is produced (e.g., single subunit per operon-encoded complex). Scenario 2: relatively little expression is required (no significantly different codon bias). Scenario 3: a hypothetical case in which high translation initiation results in ribosome jamming because the elongation rate is not optimized. Scenario 4: relatively much protein is produced (e.g., multiple subunits per operon-encoded complex); high translation initiation (low RNA fold), elevated but not maximal elongation rate (codon adaptation), elevated but not maximal ribosome densities (experimental profiles), and high protein yield (experimental protein complex stoichiometry values). Scenario 5 will also lead to high protein yields, but the experimentally detected elevated ribosome densities indicate that (at least under the tested conditions) elongation rates are not maximal. In conclusion, scenario 4 (red box) appears to most closely approach the in vivo situation of translation-controlled overexpression.

initiation rates. Indeed, in several instances, we found significant correlations between subunit stoichiometry and the mRNA folding energy in the RBS region of the analyzed genes (Figures 1, S2, and S3; Table S2). However, we obtained a relatively high number (~20%) of false positives, i.e., genes with the highest folding energy that did not code for the most abundant subunit. This might be a consequence of difficulties in predicting the correct mRNA structure, since the setting of the RNA fold analysis (i.e., the selected sequence window to be analyzed) has a strong influence on the outcome of the structure prediction (Lange et al., 2012). Instead, subunit abundance showed the strongest correlation with the fraction of optimal codons in a gene (Figures 1, S2, and S3; Table S2). Thus, codon optimization is the most reliable in silico indicator of subunit stoichiometry of operon-encoded complexes.

Complementary to the in silico analyses, a meta-analysis of experimentally determined ribosome density profiles confirms the widespread occurrence of differential translation that explains operon-encoded protein expression. The observation of increased ribosome densities implies that translation initiation is significantly increased for genes encoding highly abundant proteins in the selected operons. The correlation between codon adaptation and subunit stoichiometry reflects enhanced translation elongation rates in the cistrons coding for abundant subunits (Gingold and Pilpel, 2011; Figure 3). The increased rate of translation elongation could contribute to the avoidance of ribosome crowding (Tuller et al., 2010).

Our findings suggest that internal translation initiation is essential for differential translation of cistrons at any position within polycistronic transcripts (Figure 4; Movies S1 and S2). De novo initiation of translation at intercistronic RBSs was previously deduced from analyses of some bacterial operons, and translational coupling was demonstrated only for a subset of the genes in an operon (McCarthy, 1990; Oppenheim and Yanofsky, 1980). In the present study, we provide in silico and in vivo data that indicate the frequent occurrence of de novo intercistronic initiation of translation (Figure 1; Table S2). Moreover, we demonstrate that uncoupled translation of polycistronic messengers allows for differential translation (Figures 2 and S6). This initial experimental support for the model can serve as a prelude to wider experimental testing using both natural operon sequences and synthetic biology approaches.

It should be emphasized that in addition to the ten widespread operons analyzed in this study, many more (if not all) prokaryotic protein complexes with uneven stoichiometry are likely to rely on differential translation for tuning of their protein levels. Similarly, operons that encode enzymes of metabolic pathways might employ differential translation in case the enzymes are required in uneven quantities. Overall, we conclude that modulation of translation efficiency is a universal mode of control in bacteria and archaea that allows for differential protein expression of operon-encoded multisubunit complexes with uneven stoichiometry. This fundamental principle can be applied for prediction of the ratios between protein subunits of uncharacterized complexes (e.g., CRISPR-Cas complexes) as well as for the design of synthetic operons.

EXPERIMENTAL PROCEDURES

Transcription Analysis

The expression of genes and their division into transcriptional units were analyzed by a combination of high-throughput complementary DNA (cDNA) sequencing data (for *E. coli*, *Sulfolobus solfataricus*, and *Pseudomonas aeruginosa*), TSS maps (for *E. coli*, *S. solfataricus*, *P. aeruginosa*, and *Listeria monocytogenes*), and tiling array data (*L. monocytogenes*) (Cho et al., 2009; Mendoza-Vargas et al., 2009; Toledo-Arana et al., 2009; Wurtzel et al., 2010). See the Extended Experimental Procedures for more detail.

Translation Analysis

Genes constituting 12 different operons encoding multisubunit complexes of known stoichiometry were obtained from the NCBI database. For each gene, the codon bias was analyzed using the CAI (Sharp and Li, 1987), percentages of optimal codons (F) and codon co-occurrence (Co). Positions around the start codon were used to calculate the RNA folding energy (E) or the RNA hybridization energy between the SD/anti-SD sequences. All obtained values were averaged within each operon. The operon mean values were subtracted from the individual gene bias values, producing shift values (ΔG, ΔF, ΔCAI, ΔCo, and ΔE) for each gene, which were finally averaged across gene families. See the Extended Experimental Procedures for more detail.

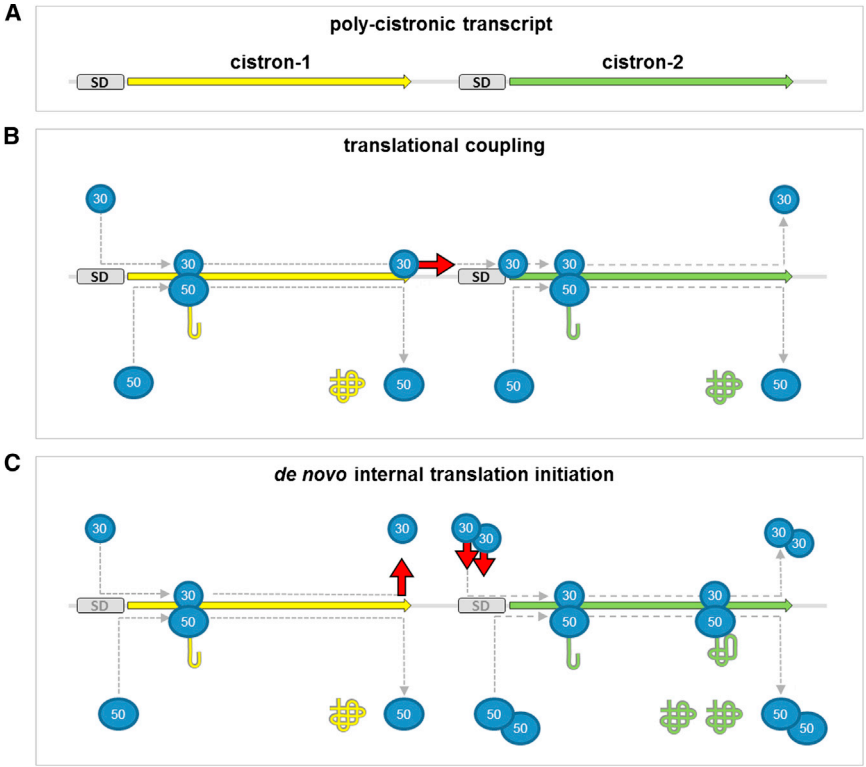

Figure 4. Model for Translation of Polycistronic Messengers in Prokaryotes

(A) Transcript consisting of two cistrons, each preceded by an SD sequence.

(B) Translational coupling where the 30S ribosomal subunit remains associated (red arrow) after termination, and 50S joins for reinitiation, resulting in stoichiometric output from both cistrons (Movie S1).

(C) De novo internal recruitment (red arrows) of both 50S and 30S subunits allows for differential translation initiation rates between cistrons. Depending on the translation elongation rate (codon bias) of each cistron, this may result in different ribosome density profiles (Movie S2); different types of broken arrows reflect different elongation rates (yellow cistron, low; green cistron, high); only in case of concommitant elevated initiation rate (panel-c, not in panel-b) this will result in increased protein production. Yellow and green cistrons have low and high translation efficiency, respectively (Figure 2).

Plasmid Constructs and Protein Detection

We synthesized operons containing two genes with synonymous mutations encoding identical GFP polypeptides and expressed in *E. coli* BL21(DE3). Two hours after isopropyl β-D-1-thiogalactopyranoside (IPTG) induction, cells were harvested and cell lysates were analyzed by western blot. See the Extended Experimental Procedures for more detail.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Extended Experimental Procedures, six figures, two tables, and two movies and can be found with this article online at http://dx.doi.org/10.1016/j.celrep.2013.07.049.

ACKNOWLEDGMENTS

We thank Jonathan Weissman and Gene-Wei Li (UCSF) for providing ribosome profiling data and stimulating discussions. This project was funded by the Netherlands Organization for Scientific Research (NWO) via an ALW TOP grant (854.10.003) to J.v.d.O. and an ALW Vidi grant (864.11.005) to S.J.J.B. T.E.F.Q. was supported by a fellowship from the Ministère de l'Enseignement Supérieur et de la Recherche de France. Y.I.W., W.R., K.S.M., and E.V.K. were supported by the U.S. Department of Health and Human Services intramural program (National Library of Medicine, NIH). O.W. was supported by an Azrieli PhD fellowship, and R.S. was supported by an ERC-StG grant (260432) and the Israeli Science Foundation (grant ISF-1303/12). A.C.F. and E.G.H.W. were supported by grants from the Swedish Research Council, including the Uppsala RNA Research Centre.

Received: December 20, 2012
Revised: July 15, 2013
Accepted: July 31, 2013
Published: September 5, 2013

REFERENCES

Abrahams, J.P., Leslie, A.G., Lutter, R., and Walker, J.E. (1994). Structure at 2.8 A resolution of F1-ATPase from bovine heart mitochondria. Nature *370*, 621–628.

Beyenbach, K.W., and Wieczorek, H. (2006). The V-type H+ ATPase: molecular structure and function, physiological roles and regulation. J. Exp. Biol. *209*, 577–589.

Brenner, S., Jacob, F., and Meselson, M. (1961). An unstable intermediate carrying information from genes to ribosomes for protein synthesis. Nature *190*, 576–581.

Cannarozzi, G., Schraudolph, N.N., Faty, M., von Rohr, P., Friberg, M.T., Roth, A.C., Gonnet, P., Gonnet, G., and Barral, Y. (2010). A role for codon order in translation dynamics. Cell *141*, 355–367.

Cho, B.K., Zengler, K., Qiu, Y., Park, Y.S., Knight, E.M., Barrett, C.L., Gao, Y., and Palsson, B.O. (2009). The transcription unit architecture of the Escherichia coli genome. Nat. Biotechnol. *27*, 1043–1049.

Drummond, D.A., and Wilke, C.O. (2008). Mistranslation-induced protein misfolding as a dominant constraint on coding-sequence evolution. Cell *134*, 341–352.

Dunkle, J.A., Wang, L., Feldman, M.B., Pulk, A., Chen, V.B., Kapral, G.J., Noeske, J., Richardson, J.S., Blanchard, S.C., and Cate, J.H. (2011). Structures of the bacterial ribosome in classical and hybrid states of tRNA binding. Science *332*, 981–984.

Efremov, R.G., Baradaran, R., and Sazanov, L.A. (2010). The architecture of respiratory complex I. Nature *465*, 441–445.

Errington, J. (2003). Dynamic proteins and a cytoskeleton in bacteria. Nat. Cell Biol. *5*, 175–178.

Forster, A.C. (2012). Synthetic biology challenges long-held hypotheses in translation, codon bias and transcription. Biotechnol. J. *7*, 835–845.

Fredrick, K., and Ibba, M. (2010). How the sequence of a gene can tune its translation. Cell *141*, 227–229.

French, S.L., Santangelo, T.J., Beyer, A.L., and Reeve, J.N. (2007). Transcription and translation are coupled in Archaea. Mol. Biol. Evol. *24*, 893–895.

Ghosh, A., and Albers, S.V. (2011). Assembly and function of the archaeal flagellum. Biochem. Soc. Trans. *39*, 64–69.

Gingold, H., and Pilpel, Y. (2011). Determinants of translation efficiency and accuracy. Mol. Syst. Biol. *7*, 481.

Goldberger, R.F., Deeley, R.G., and Mullinix, K.P. (1976). Regulation of gene expression in prokaryotic organisms. Adv. Genet. *18*, 1–67.

Grundy, F.J., and Henkin, T.M. (2006). From ribosome to riboswitch: control of gene expression in bacteria by RNA structural rearrangements. Crit. Rev. Biochem. Mol. Biol. *41*, 329–338.

Gur, E., Biran, D., and Ron, E.Z. (2011). Regulated proteolysis in Gram-negative bacteria—how and when? Nat. Rev. Microbiol. *9*, 839–848.

Ikemura, T. (1981). Correlation between the abundance of Escherichia coli transfer RNAs and the occurrence of the respective codons in its protein genes. J. Mol. Biol. *146*, 1–21.

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.

Jacob, F., and Monod, J. (1961). Genetic regulatory mechanisms in the synthesis of proteins. J. Mol. Biol. *3*, 318–356.

Jakob, M., Kaiser, S., Gutensohn, M., Hanner, P., and Klösgen, R.B. (2009). Tat subunit stoichiometry in Arabidopsis thaliana challenges the proposed function of TatA as the translocation pore. Biochim. Biophys. Acta *1793*, 388–394.

Johnson, T.L., Abendroth, J., Hol, W.G., and Sandkvist, M. (2006). Type II secretion: from structure to function. FEMS Microbiol. Lett. *255*, 175–186.

Jore, M.M., Lundgren, M., van Duijn, E., Bultema, J.B., Westra, E.R., Waghmare, S.P., Wiedenheft, B., Pul, U., Wurm, R., Wagner, R., et al. (2011). Structural basis for CRISPR RNA-guided DNA recognition by Cascade. Nat. Struct. Mol. Biol. *18*, 529–536.

Kaberdin, V.R., and Bläsi, U. (2006). Translation initiation and the fate of bacterial mRNAs. FEMS Microbiol. Rev. *30*, 967–979.

Koide, T., Reiss, D.J., Bare, J.C., Pang, W.L., Facciotti, M.T., Schmid, A.K., Pan, M., Marzolf, B., Van, P.T., Lo, F.Y., et al. (2009). Prevalence of transcription promoters within archaeal operons and coding sequences. Mol. Syst. Biol. *5*, 285.

Koonin, E.V. (2009). Evolution of genome architecture. Int. J. Biochem. Cell Biol. *41*, 298–306.

Kudla, G., Murray, A.W., Tollervey, D., and Plotkin, J.B. (2009). Coding-sequence determinants of gene expression in Escherichia coli. Science *324*, 255–258.

Lange, S.J., Maticzka, D., Möhl, M., Gagnon, J.N., Brown, C.M., and Backofen, R. (2012). Global or local? Predicting secondary structure and accessibility in mRNAs. Nucleic Acids Res. *40*, 5215–5226.

Li, Y., and Altman, S. (2004). Polarity effects in the lactose operon of Escherichia coli. J. Mol. Biol. *339*, 31–39.

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.

Marzi, S., Myasnikov, A.G., Serganov, A., Ehresmann, C., Romby, P., Yusupov, M., and Klaholz, B.P. (2007). Structured mRNAs regulate translation initiation by binding to the platform of the ribosome. Cell *130*, 1019–1031.

McCarthy, J.E. (1990). Post-transcriptional control in the polycistronic operon environment: studies of the atp operon of Escherichia coli. Mol. Microbiol. *4*, 1233–1240.

Mendoza-Vargas, A., Olvera, L., Olvera, M., Grande, R., Vega-Alvarado, L., Taboada, B., Jimenez-Jacinto, V., Salgado, H., Juárez, K., Contreras-Moreira, B., et al. (2009). Genome-wide identification of transcription start sites, promoters and transcription factor binding sites in E. coli. PLoS ONE *4*, e7526.

Oppenheim, D.S., and Yanofsky, C. (1980). Translational coupling during expression of the tryptophan operon of Escherichia coli. Genetics *95*, 785–795.

Shao, Z.Q., Zhang, Y.M., Feng, X.Y., Wang, B., and Chen, J.Q. (2012). Synonymous codon ordering: a subtle but prevalent strategy of bacteria to improve translational efficiency. PLoS ONE *7*, e33547.

Sharp, P.M., and Li, W.H. (1987). The codon Adaptation Index—a measure of directional synonymous codon usage bias, and its potential applications. Nucleic Acids Res. *15*, 1281–1295.

Stenström, C.M., Jin, H., Major, L.L., Tate, W.P., and Isaksson, L.A. (2001). Codon bias at the 3′-side of the initiation codon is correlated with translation initiation efficiency in Escherichia coli. Gene *263*, 273–284.

Timmermans, J., and Van Melderen, L. (2010). Post-transcriptional global regulation by CsrA in bacteria. Cell. Mol. Life Sci. *67*, 2897–2908.

Toledo-Arana, A., Dussurget, O., Nikitas, G., Sesto, N., Guet-Revillet, H., Balestrino, D., Loh, E., Gripenland, J., Tiensuu, T., Vaitkevicius, K., et al. (2009). The Listeria transcriptional landscape from saprophytism to virulence. Nature *459*, 950–956.

Tuller, T., Carmi, A., Vestsigian, K., Navon, S., Dorfan, Y., Zaborske, J., Pan, T., Dahan, O., Furman, I., and Pilpel, Y. (2010). An evolutionarily conserved mechanism for controlling the efficiency of protein translation. Cell *141*, 344–354.

Vellanoweth, R.L., and Rabinowitz, J.C. (1992). The influence of ribosome-binding-site elements on translational efficiency in Bacillus subtilis and Escherichia coli in vivo. Mol. Microbiol. *6*, 1105–1114.

Welch, M., Govindarajan, S., Ness, J.E., Villalobos, A., Gurney, A., Minshull, J., and Gustafsson, C. (2009). Design parameters to control synthetic gene expression in Escherichia coli. PLoS ONE *4*, e7002.

Wiedenheft, B., van Duijn, E., Bultema, J.B., Waghmare, S.P., Zhou, K., Barendregt, A., Westphal, W., Heck, A.J., Boekema, E.J., Dickman, M.J., and Doudna, J.A. (2011). RNA-guided complex from a bacterial immune system enhances target recognition through seed sequence interactions. Proc. Natl. Acad. Sci. USA *108*, 10092–10097.

Wurtzel, O., Sapra, R., Chen, F., Zhu, Y., Simmons, B.A., and Sorek, R. (2010). A single-base resolution map of an archaeal transcriptome. Genome Res. *20*, 133–141.

Wurtzel, O., Sesto, N., Mellin, J.R., Karunker, I., Edelheit, S., Bécavin, C., Archambaud, C., Cossart, P., and Sorek, R. (2012a). Comparative transcriptomics of pathogenic and non-pathogenic Listeria species. Mol. Syst. Biol. *8*, 583.

Wurtzel, O., Yoder-Himes, D.R., Han, K., Dandekar, A.A., Edelheit, S., Greenberg, E.P., Sorek, R., and Lory, S. (2012b). The single-nucleotide resolution transcriptome of Pseudomonas aeruginosa grown in body temperature. PLoS Pathog. *8*, e1002945.