ACS SyntheticBiology

Research Article

pubs.acs.org/synthbio

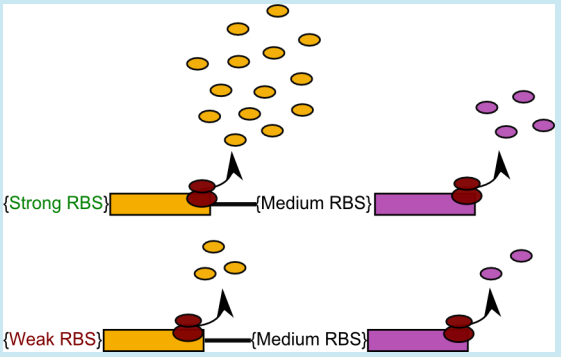
Quantifying Translational Coupling in *E. coli* Synthetic Operons Using RBS Modulation and Fluorescent Reporters

Ayelet Levin-Karp, Uri Barenholz, Tasneem Bareia, Michal Dayagi, Lior Zelcbuch, Niv Antonovsky, Elad Noor, and Ron Milo*

Department of Plant Sciences, Weizmann Institute of Science, Rehovot 76100, Israel

Supporting Information

ABSTRACT: Translational coupling is the interdependence of translation efficiency of neighboring genes encoded within an operon. The degree of coupling may be quantified by measuring how the translation rate of a gene is modulated by the translation rate of its upstream gene. Translational coupling was observed in prokaryotic operons several decades ago, but the quantitative range of modulation translational coupling leads to and the factors governing this modulation were only partially characterized. In this study, we systematically quantify and characterize translational coupling in *E. coli* synthetic operons using a library of plasmids carrying fluorescent reporter genes that are controlled by a set of different ribosome binding site (RBS) sequences. The downstream gene expression level is found to be enhanced by the upstream gene expression via translational coupling with the enhancement level varying from almost no coupling to over 10-fold depending on the upstream gene's sequence. Additionally, we find that the level of translational coupling in our system is similar between the second and third locations in the operon. The coupling depends on the distance between the stop codon of the upstream gene and the start codon of the downstream gene. This study is the first to systematically and quantitatively characterize translational coupling in a synthetic *E. coli* operon. Our analysis will be useful in accurate manipulation of gene expression in synthetic biology and serves as a step toward understanding the mechanisms involved in translational expression modulation.



KEYWORDS: *translational coupling, ribosome binding site, translation regulation, synthetic operon, synthetic biology*

The ability to reliably predict quantitative gene expression[1,2] at various contexts is essential for many synthetic biology applications.[3,4] Better understanding of the biological principles that govern gene expression will help build improved, more predictable genetic circuits. Elements controlling the intracellular abundance of protein level include, among others, the promoter that controls the transcription level, the ribosome binding site (RBS) controlling the translation initiation rate,[5] the relative stability of transcripts,[6] and the mRNA secondary structure.[7]

Additionally, the gene upstream and downstream genetic context greatly affects its expression level.[8] Genes that are involved in the same biochemical or physiological pathway are often grouped to an operon, controlling transcription of the genes to a single polycistronic mRNA. The operon is thought to reduce variability in protein level of genes by reducing transcriptional noise,[9,10] while gene location in the operon was shown to also strongly affect genes expression level via translational and mRNA stability effects.[11,12]

Translational coupling is defined as the interdependence of translation efficiency of neighboring genes encoded by the same polycistronic mRNA.[13,14] Specifically, it was shown that the translation level of a gene is increased by an enhanced translation of the gene preceding it.[15] Translational coupling was observed in various *E. coli* operons, as well as in other prokaryotes and bacteriophages.[16−19]

Given some overall expression level of a polycistronic transcript dictated by the promoter and various other factors, translational coupling can help regulate relative protein level of genes in the operon. For instance, Løvdok et al.[13] show that chemotaxis efficiency is enhanced when the chemotactic genes are coexpressed and translationally coupled. Translational coupling is suggested to be part of the robustness mechanism essential for correct chemotactic behavior.

The degree of coupling between genes varies. At the extreme, a downstream gene translation is not detected unless the upstream gene is translated. This can occur when there is a strong inhibiting secondary structure formed on the second gene's translation initiation region or when the translation initiation region sequence is too weak to initiate *de novo* translation.[20,21] More commonly, it is qualitatively observed that both genes can be translated independently, yet the second gene translation is enhanced upon translation of the first gene.[13,22]

Received: January 16, 2013
Published: April 16, 2013

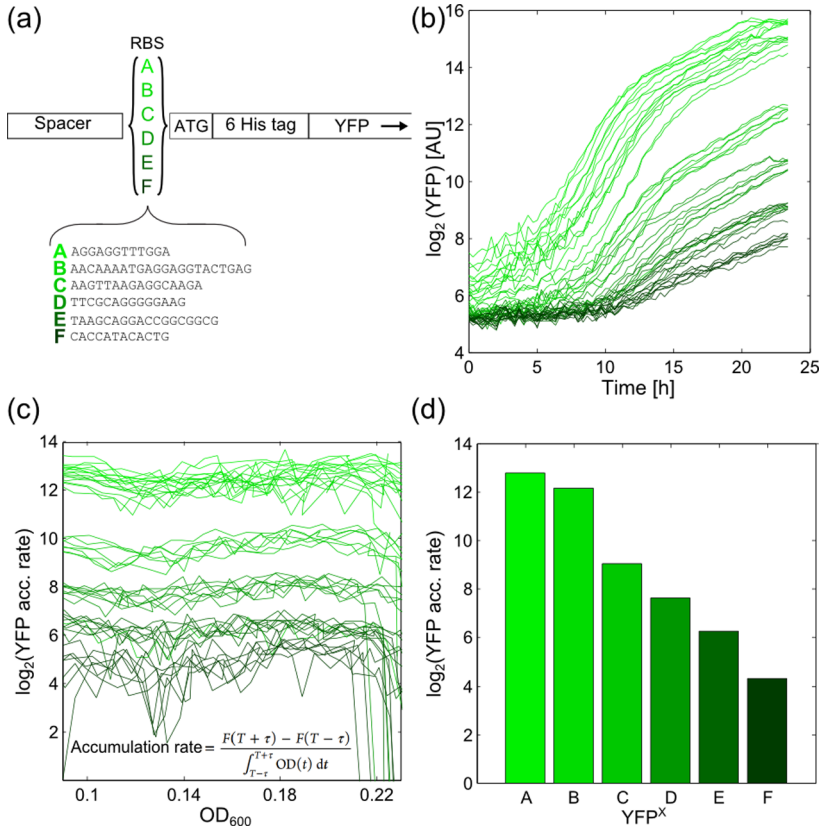


Figure 1. Construct structure for manipulating expression levels using ribosome binding sites. (a) A small set of RBS sequences was designed to span several orders of magnitude of protein expression. The RBS set was composed of six precharacterized RBS sequences (denoted in a descending order of expression strength A to F), which were paired to the gene of interest. Additionally, to minimize sequence-dependent effects, an insulator sequence of 19 bp was inserted upstream of the RBS sequence, and a 21 bp His tag sequence was inserted downstream of it. (b) YFP raw measurements (excitation at 520 nm and emission measurement at 555 nm) of growing bacteria transformed with pTac plasmid, each carrying one of the six library variants previously described. (c) YFP's accumulation rate (acc) for the different library variants. Accumulation rate ($A(T)$) was calculated by the following equation $A(T) = (F(T + \tau) - F(T - \tau))/(\int_{T-\tau}^{T+\tau} \mathrm{OD}(t)\, \mathrm{d}t)$; here F is the fluorescence measured by the detector, OD600 value is measured using a multiwell plate reader (after media subtraction). OD600 values should be multiplied by a factor of ∼3 for comparison to standard 1 cm path length optical density measurements. (d) Mean protein accumulation rate was calculated by averaging over 5 measurements around OD_{600} of 0.2 for each sample.

Two main models have been proposed to explain the mechanism of translational coupling. The first mechanism suggests that ribosomes are recruited from the stop codon of the upstream gene to the start codon of the downstream gene.[23] When the upstream gene is highly translated, more ribosomes are recruited resulting in enhanced translation of the downstream gene. Recruitment can occur by direct initiation on the downstream gene without the ribosome falling off the mRNA or by terminating translation and increasing the local concentration of ribosomal subunits next to the second gene's RBS. Research on translational coupling in the galactose operon supported such a mechanism.[14]

The second mechanism is based on the ribosome's ability to unfold mRNA secondary structures. The 70S ribosome (consisting the 30S and 50S subunits) can unfold mRNA structures and thus support translation elongation. When the 70S ribosome reaches a stop codon, it dissociates into 30S and 50S subunits. The 30S ribosome subunit has the ability to bind mRNA, scan across it, and initiate translation, but its ability to unfold structures is limited. When the first gene is actively translated, the 70S ribosome disrupts inhibitory mRNA structures in the downstream gene translation initiation region, thus allowing its 30S subunit to bind the downstream gene mRNA and initiate translation.[20,24] This model is supported by a study on the atp operon, which codes for the subunits of the membrane-associated proton-translocating ATP synthase. A secondary structure analysis of the *atp* operon revealed a stem loop structure at the end of the upstream gene (*atp*H) that blocks the translation initiation region of the downstream gene (*atp*A). Only when the first gene is translated does the inhibiting secondary structure open up and allow translation of the second gene. Mutating the downstream gene's translation initiation sequence such that the inhibiting structure is not formed results in elevated expression. More supportive evidence for the secondary structure model are reports of operons where the coupling is observed to depend on the upstream gene sequence.[25,26] Notably, the two mechanisms described above for translational coupling do not contradict each other and can operate concurrently to give rise to the overall translation coupling modulation observed.

While translational coupling was previously reported for several different operons, most studies focused on describing the coupling in native operons cloned on expression plasmids, attempting to understand the underlying coupling mechanism. A systematic quantification of the coupling is still lacking, and the factors affecting it were only very limitedly explored. In this study, we attempt to systematically characterize the translational coupling effect in synthetic operons, which are currently

dx.doi.org/10.1021/sb400002n | *ACS Synth. Biol.* 2013, 2, 327−336

widely used in synthetic biology and metabolic engineering efforts. Emphasis is put on trying to describe the effect quantitatively rather than qualitatively. We measure the range of fold change modulation that arises due to translational coupling and characterize how translational coupling is affected by changes in gene sequence, the relative position in the operon and the translational distance between genes.

RESULTS AND DISCUSSION

In order to quantitatively measure the translational coupling effect, we constructed a synthetic bicystronic operon library, containing YFP followed by RFP (Y−R library). By using six short RBS sequences just upstream of a gene, denoted RBS^A for strongest RBS to RBS^F for weakest RBS, we could manipulate protein expression by ∼20−200-fold (Figure 1, Supplementary Table S1). In each operon variant, YFP was controlled by one of six RBS sequences (YFP^{A-F}) while RFP was controlled by one of three RBS sequences (RFP^A, RFP^C, RFP^E).[27] In our synthetic operon system the genes are separated by ∼50 bp and which include an identical separating sequence that can be thought to serve as an insulator. The translation initiation region of the downstream gene does not overlap with the upstream gene (for a more detailed description of the construct see Figure 1 and Supplementary Tables S2 and S3).

At each time point where the fluorescence intensity was measured, the protein accumulation rate was calculated using the change in the fluorescence level divided by the total producing biomass approximated by the OD as described in detail in the Methods section (see also Figure 1). Since the degradation rate of fluorescent proteins is assumed to be slow with respect to the cell doubling time, the accumulation rate is close to the protein production rate and thus a good proxy for a gene activity level when bacteria are growing exponentially. We chose to use the protein production rate rather than the protein level (fluorescence level divided by OD) to analyze our data, as translational coupling models predict that the activity level of the upstream gene affects the activity level of the downstream gene. Therefore we believe protein accumulation rate is a more adequate way to describe translational coupling dynamics (it seems rather unlikely that the number of proteins from the upstream gene in the cell, the protein level itself, will induce the elevation in the protein level of the downstream gene). Notably, we do not see any qualitative difference when analyzing the data using protein level or protein accumulation rate as shown in Supplementary Figure S2.

The Y−R library variants were transformed to *E. coli K12* MG1655, and their fluorescence intensity was measured in both RFP and YFP's channels. The results reveal a coupling effect; the upstream gene (YFP) protein accumulation rate is affected only by its RBS-induced translation rate, while the protein accumulation rate of the downstream gene (RFP) depends both on its RBS sequence, as well as on the expression level of the upstream gene (YFP) (Figure 2). Modulation from weak to strong RBS in the upstream gene (∼200-fold) enhances the downstream gene expression level by ∼5-fold. Notably, the magnitude of the coupling effect was similar regardless of the downstream gene RBS sequence, as seen by the similarity in the blue (RBS^A), red (RBS^C), and green (RBS^E) trends of Figure 2.

In order to find characteristic ranges for the change in expression via translational coupling when an upstream gene is modulated between a strong and a weak RBS, we repeated the analysis for various gene and operon arrangements. As shown in Figure 3 (orange bars), the coupling modulation is

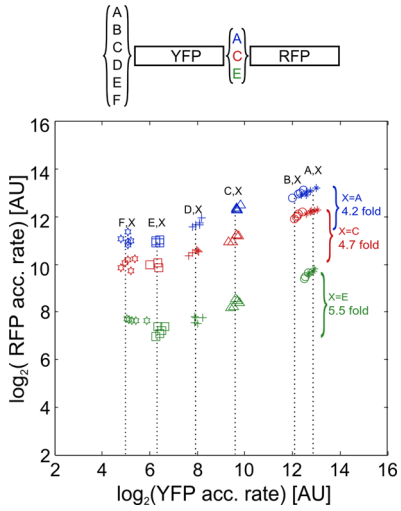


Figure 2. Quantitative characterization of translational coupling using a fluorescent reporter system. YFP and RFP protein accumulation rate of the Y−R library, which is composed of YFP followed by RFP. At each operon variant, YFP and RFP were controlled by different RBS sequences (their set denoted as X, such that C, X indicates that the first gene has RBS C and the second gene is modulated by different RBS sequences from the library as indicated for each point by its corresponding color). RFP was controlled by one of three RBS sequences, denoted using colors (blue for the strong RBS^A, red for the medium RBS^C, and green for the weak RBS^E), and YFP was controlled by one of the six RBS's corresponding to shapes (asterisk for RBS^A, circle for RBS^B, triangle for RBS^C, cross for RBS^D, square for RBS^E, and star for RBS^F). Fold change expression of the downstream RFP induced by changing from the weakest to the strongest RBS variant of the upstream YFP is indicated on the right-hand side. We find an over 4-fold increase in downstream expression as the upstream gene in the operon is modulated by about 100-fold.

characteristically about 5-fold and varies between ∼1.3-fold (almost no coupling) to ∼15-fold (strong coupling) depending on the genes at the two positions (see figure caption for full description). We find that the coupling effect is common in synthetic operons, yet the exact degree of enhancement is context-dependent (Figure 3).

To test if the observed coupling is mainly a translational or a transcriptional effect, RT-qPCR measurements were performed, measuring the mRNA level of a gene when its upstream gene is controlled by a strong or a weak RBS. The RT-qPCR results (Figure 3, purple bars) show that while the fluorescence level of the protein is modulated in most cases by ∼5−15-fold, the mRNA level change was not significant (except for one case where it changed by ∼4-fold). This suggests that the coupling effect we observe is not mainly explained by transcription modulation or mRNA stability effects and is dominated by a translational effect.

Translational Coupling Depends on the Upstream Gene Specific Sequence and Not Only on Expression Level. To better understand the impact of gene sequence on translational coupling, we next aimed to measure translational coupling in an inverted operon where YFP and RFP locations are switched. In this library (referred to as R−Y) RFP is upstream and modulated by RBSs A−F, and YFP is downstream controlled by RBSs A, C, or E (Figure 4a).

The upstream gene (RFP) protein accumulation rate was affected only by its own RBS and not by the protein accumulation rate of the downstream gene (YFP). This is

dx.doi.org/10.1021/sb400002n | *ACS Synth. Biol.* 2013, 2, 327−336

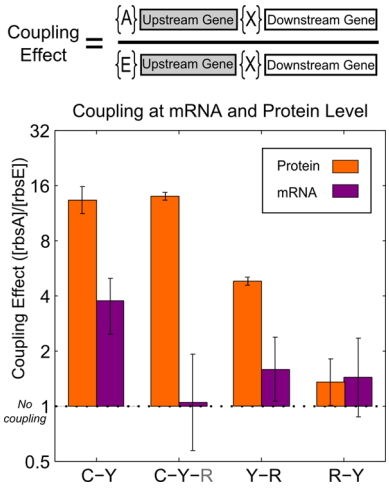


Figure 3. Translational coupling strength is context-dependent and cannot be explained by changes in mRNA levels alone. Translational coupling enhancement for various gene and operon arrangements (C–Y: CFP followed by YFP^C; C–Y–R: CFP followed by YFP^E followed by RFP^E (measuring YFP); Y–R: YFP followed by RFP^C; R–Y: RFP followed by YFP^C). Coupling enhancement of the protein level (orange bars) was calculated by dividing the expression level of the downstream gene, when the upstream gene is under a strong RBS (RBS^A), with the expression level of the downstream gene when the upstream gene is under a weak RBS (RBS^E) as shown in inset. Coupling enhancement at mRNA level (purple bars) was calculated similarly using RT-qPCR measurements of the downstream gene. We find that the coupling effect is not explained by changes in mRNA levels.

unlike YFP's effect on RFP's protein accumulation rate in the reversed Y–R operon when YFP is upstream to RFP (Figure 2). Thus, the modulation we observe is seen to be directional, which is consistent with the definition of translational coupling rather than some other coupling mechanism (Figure 4a).

When examining YFP's expression in the R–Y operon, very little translational coupling is observed. Maximum translation enhancement effect of ~2-fold is measured when YFP is controlled by a weak RBS (YFP^E), whereas when YFP is controlled by a strong RBS (YFP^A), no coupling effect is measured at all (Figure 4a). The smaller coupling effect may be due to some inherent properties of the downstream gene (YFP) or due to the change in the upstream gene sequence (RFP). To distinguish between these two options another library was constructed, this time using CFP instead of RFP as the upstream gene in front of the downstream YFP (referred to as C–Y library). CFP and YFP fluorescence measurements revealed a strong translational coupling effect with a maximal translation enhancement of over 10-fold (low expression levels make a more accurate estimate difficult) (Figure 4b). Thus, having CFP instead of RFP upstream to YFP re-establishes translational coupling, exemplifying that translational coupling depends on the upstream gene sequence.

Translational Coupling in Our Synthetic Operon Shows Similar Coupling Enhancement Values between the Second and Third Gene Location. Native and synthetic operons often contain more than two genes. We were interested in how translational coupling is affected by the larger operon context. To answer this question, we first wanted to see the effect of adding a downstream gene. We thus built a plasmid library combining three reporter genes CFP-YFP-RFP (referred to as C–Y–R library). Each plasmid was controlled by one of three RBS sequences (A, C, E). Figure 5 compares the CFP and YFP measurements of two constructs, the C–Y–R library and the C–Y library. Both libraries show similar protein accumulation rates for CFP and YFP. Thus, translational coupling in our system is comparable when RFP is added downstream to the coupled gene.

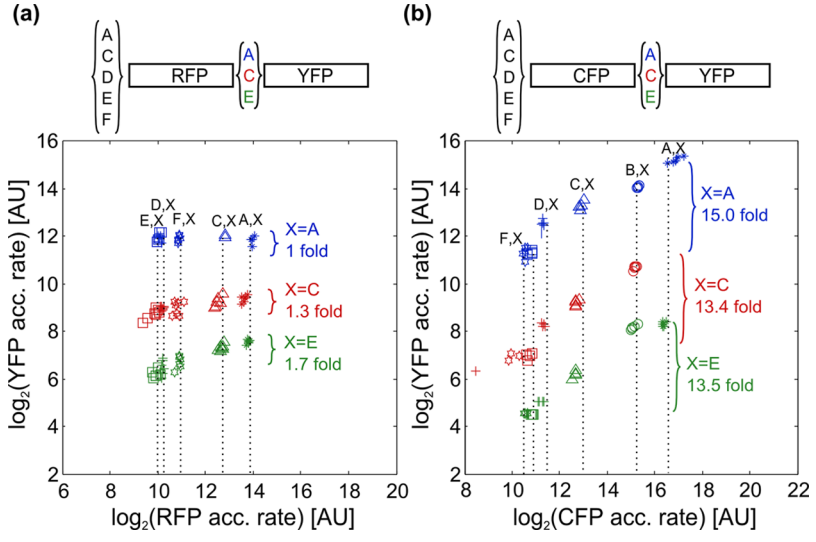


Figure 4. Translational coupling depends on the specific sequence of the upstream gene and not only on its expression level. (a) YFP and RFP protein accumulation rate measurements of an operon composed of RFP followed by YFP. At each operon variant RFP and YFP are controlled by different RBS sequences. We find very weak coupling in this operon arrangement. RFP RBS sequence corresponds to shapes (asterisk for RBS^A, triangle for RBS^C, cross for RBS^D, square for RBS^E, and star for RBS^F), while RBS controlling YFP are denoted using colors (blue for RBS^A, red for RBS^C, and green for RBS^E). (b) Schematic diagram and fluorescent measurements of the C–Y operon, which is composed of CFP (instead of RFP) followed by YFP. We find much stronger coupling in this operon arrangement. The CFP controlling RBS sequences are denoted by shapes (asterisk for RBS^A, circle for RBS^B, triangle for RBS^C, cross for RBS^D, square for RBS^E, and star for RBS^F), while YFP controlling RBS sequences are denoted using colors (blue for RBS^A, red for RBS^C, and green for RBS^E). Fold change induced by coupling for each YFP variant is indicated.

dx.doi.org/10.1021/sb400002n | ACS Synth. Biol. 2013, 2, 327–336

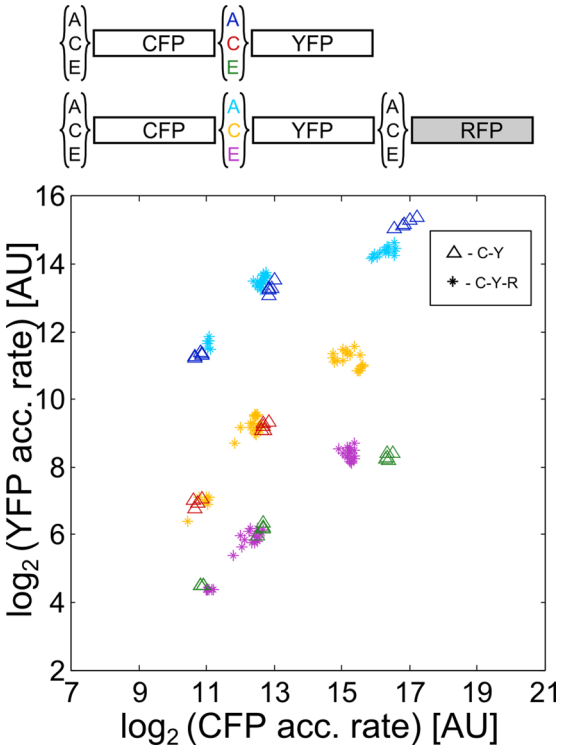


Figure 5. Translational coupling in our synthetic operon is robust to the addition of a downstream gene. CFP and YFP protein accumulation rate measurements in a CFP-YFP operon versus the CFP-YFP-RFP operon. Each gene is controlled by one of three RBS sequences. We find that the addition of the downstream third RFP gene did not alter the expression of the first two genes. YFP RBS's are denoted using colors (for C−Y library: blue for RBS^A, red for RBS^C, and green for RBS^E; for C−Y−R library: azure for RBS^A, yellow for RBS^C, and purple for RBS^E). Additionally, the C−Y library is denoted with triangles, and the C−Y−R library is denoted with asterisks. In some constructs when CFP was controlled by the weak RBS^E, the values of CFP were too low to be discerned above the background and thus the accumulation rate could not be calculated.

We next aimed to examine if the coupling effect is sensitive to the gene location in the operon (relative to the promoter). For that we compared the protein accumulation rate of two genes (YFP and RFP) when located at the first and second locations in the operon (Y−R library) (Figure 6a) to their protein accumulation rates when placed at the second and third locations on the operon (C−Y−R library) (Figure 6b). To isolate the coupling effect of YFP on RFP in the C−Y−R operon from the effect of CFP, we focused only on the results where CFP was controlled by RBS^C (denoted CFP^C). The intensity of modulation of RFP by YFP when placed as second and third is similar to their translational coupling pattern when alone on the operon (Figure 6). The same is true when the upstream CFP is controlled by RBS^A and RBS^E (see Supplementary Figure S3). Thus, translational coupling in our system is comparable between the second and third gene locations for the few cases we tested. More systematic quantification remains an open challenge.

A Cascading Translational Coupling Effect Is Evident between the First and Third Genes. As shown above, we observe translational coupling between the first and second as well as second and third gene locations in the operon. The question then arises whether this effect is limited only to adjacent genes. Using the C−Y−R tricolor reporter operon, we were able to quantify the translational effect going beyond two adjacent genes. Analyzing the data we find an enhancement of ∼5-fold in the third gene (RFP) expression level when the first gene (CFP) RBS is modulated while the second gene (YFP) RBS is kept constant (YFP^C) (Figure 7a).

Can this effect be explained by the increase in the second gene expression resulting from its coupling to the first gene, which in turn causes a rise in the third gene? In Figure 4b we show that the change in the first gene RBS sequence (CFP) modulates the second gene (YFP) protein accumulation rate by ∼15-fold (Figure 4b). This suggests that a gene's expression level could be affected not only by its closest upstream gene RBS sequence but also by a further removed upstream gene expression level via a cascade of translational coupling effects.

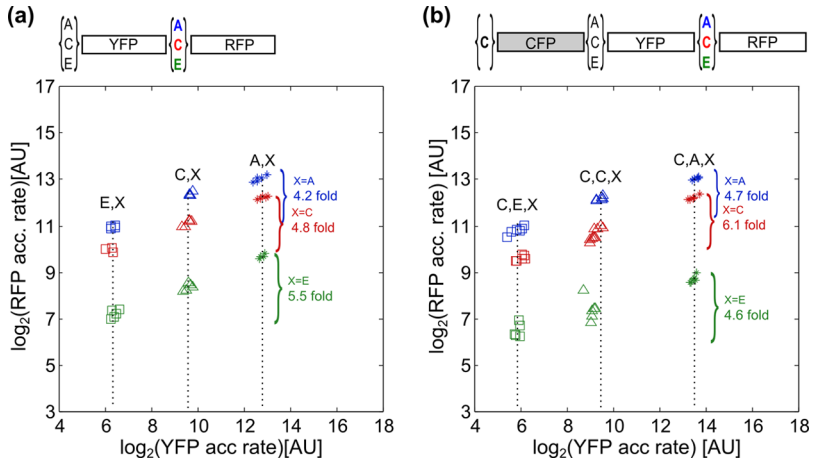


Figure 6. Translational coupling in our synthetic operon is robust to location on the operon. (a) YFP and RFP protein accumulation rate when first and second in the operon (Y−R library). At each operon variant YFP (first) and RFP (second) are controlled by one of three different RBS sequences (RBS^A, RBS^C, RBS^E). YFP's RBS sequence denoted by shapes (asterisk for RBS^A, triangle for RBS^C, and square for RBS^E), while RBS controlling RFP are denoted using colors (blue for RBS^A, red for RBS^C, and green for RBS^E). (b) YFP and RFP protein accumulation rate when second and third genes on the operon (C−Y−R library), while the first gene (CFP) RBS is kept constant (CFP^C). YFP (second) and RFP (third) are controlled by one of three different RBS sequences (RBS^A, RBS^C, RBS^E). The intensity of modulation of RFP by YFP when placed as second and third is similar to their translational coupling pattern when alone on the operon.

dx.doi.org/10.1021/sb400002n | ACS Synth. Biol. 2013, 2, 327−336

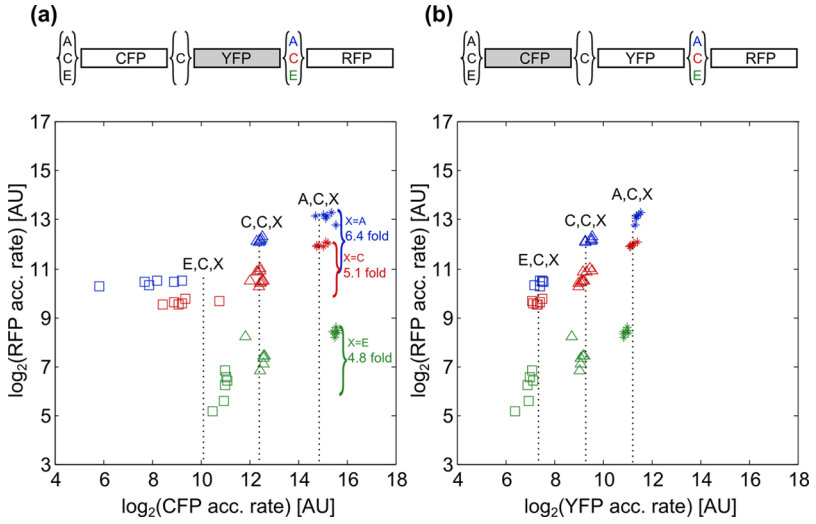


Figure 7. Cascading translational coupling effect between the first and third gene in an operon. (a) CFP and RFP accumulation rate of the third versus the first gene in the CFP-YFP-RFP operon. At each operon variant CFP (first) and RFP (third) are controlled by one of three different RBS sequences (RBS^A, RBS^C, RBS^E), while YFP (second gene) RBS is kept constant (YFP^C). CFP RBS sequence corresponds to shapes (asterisk for RBS^A, triangle for RBS^C ,and square for RBS^E), while RBSs controlling RFP are denoted using colors (blue for RBS^A, red for RBS^C, and green for RBS^E). (b) Accumulation rate of the second and third genes (YFP and RFP) from the same experiment. Fold change modulation is not shown, as it is the same as in panel a (both show RFP level of the same constructs on the *y*-axis).

The coupling effect in this case is even somewhat stronger than the one seen in the direct modulation analyzed for Y−R library (Figure 2) where a modulation of ∼200-fold in YFP was needed to achieve an ∼5-fold translational enhancement of the downstream RFP gene. Similar results are obtained when the middle gene (YFP) is controlled by RBS^A and RBS^E (Supplementary Figure S4).

Translational Coupling Depends on the Distance between the Stop Codon and Downstream Start Codon. Finally, we aimed to characterize the dependence of the coupling effect on the distance between the stop codon of the upstream gene and the start codon of the downstream gene. To accomplish this, a stop codon was inserted in the upstream gene (YFP) sequence at −150, −250, −350, and −850 bp in two of the Y−R library constructs: YFP^A-RFP^C and YFP^C-RFP^C. Fluorescence measurements of the downstream gene (RFP) show that the insertion of an early stop codon at the upstream gene significantly decreases the translational coupling effect, hence decreasing the protein accumulation rate of the downstream gene. The level of reduction seems to mostly correlate with the translational distance: For YFP^A-RFP^C the trend seems consistent with an exponential decay for the first 350 bp, which gives a decrease of ∼1.4-fold every ∼100 bp when a stop codon is inserted close to the end of the gene. This trend does not continue throughout the gene as the maximal decrease observed at −850 bp (no start codon denoted, noATG) is ∼5.5-fold. For YFP^C-RFP^C constructs the picture is similar, only that the decrease is ∼2-fold for every ∼100 bp between 20 and 250 bp (and there is a slight unexpected but repeatable increase of ∼2-fold in the coupling when moving to −300 bp). Again, it is not likely that this trend continues throughout the gene as the max decrease observed at −850 bp (noATG) is ∼8-fold. In short, the greater the distance, the stronger the reduction in the expression level of the downstream gene, i.e., the weaker is the coupling effect (Figure 8). This indicates that translational coupling depends on the translational distance between genes in a graded rather than abrupt manner.

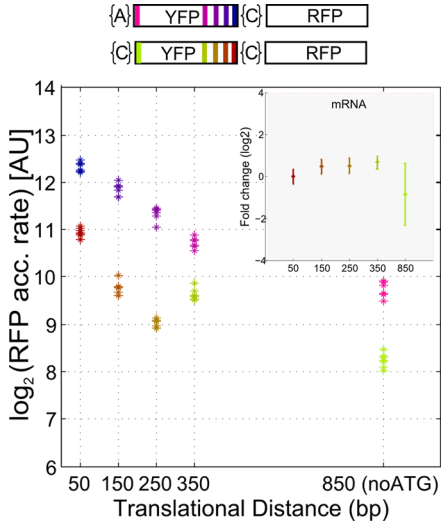


Figure 8. Translational coupling decreases continuously with the distance between the stop codon and downstream start codon. RFP's protein accumulation rate measurements of the second gene in an operon, when a stop codon is inserted in different locations in the upstream gene. Two cases are shown, one with the upstream YFP gene with the strong RBS^A (blue to pink dots, depending on stop location) and the second with the medium RBS^C (red to green dots). We find that as the stop codons alter the translational distance between the genes from 50 bp (the natural stop codon location in our constructs) to 850 bp (stop codon was inserted instead of YFP's start codon), the downstream expression continuously decreases. Constructs are denoted according to the distance (bp) between the inserted stop codon and the start codon of the downstream gene (RFP). (Inset) A RT-qPCR measurement of RFP mRNA levels for the YFP^C-RFP^C constructs showing the effect is not due to mRNA levels.

Inserting early stop codons might enhance mRNA degradation, as it stops ribosomes from translating the entire transcript and can facilitate RNases access to the mRNA. To test if the measured reduction is a result of mRNA instability, RT-qPCR measurements were performed for YFP^C-RFP^C

dx.doi.org/10.1021/sb400002n | *ACS Synth. Biol.* 2013, 2, 327−336

constructs. No significant decrease in YFP or RFP mRNA levels was found between plasmids with different translational distances (Figure 8, inset).

Surprisingly, in a construct where we deleted the ATG start site in an operon upstream gene we still observed a relative elevation in the downstream gene level (YFP^{A}-RFP^{C} versus YFP^{C}-RFP^{C}). We find that this elevation is evident at the mRNA level as well, which suggests that in this case there is a relative decrease in mRNA stability (Supplementary Figure S5).

Our results on incorporating early stop codons are in line with previous findings in a similar experiment that studied the Galactose operon.[14] In that study a decrease of about 30% with a stop codon 50 bp upstream and ∼60% decrease with a stop codon 150 bp upstream was observed. This observation is similar to the ∼2-fold decrease we observe when shifting the stop codon ∼100 bp upstream (50 versus 150 in Figure 8). Furthermore, in the Galactose operon study a decrease in coupling was observed when genes overlap (stop codon of the first gene is downstream to the second gene start codon[14]).

Conclusions and Comparison to Previous Studies on Translational Coupling. Most previous studies of translational coupling measured native *E. coli* operons that were reconstructed on plasmids. In these models the genes are usually a few nucleotides apart or even overlapping. Oppenheim and Yanofsky[26] show a great decrease in coupling between a gene pair 1 bp apart compared to a gene pair 11 bp apart.

We find coupling at larger distances, for example, about an order of magnitude of translational coupling enhancement found for the CFP-YFP gene pair, when the genes are ∼50 bp apart. Such a substantial effect can be important when fine-tuning gene expression in synthetic operons.

A large variation in translational coupling enhancement was measured in our system, ranging from almost no effect (RFP-YFP^{A}) to a coupling enhancement of ∼15-fold in gene expression (CFP-YFP^{E}) when the upstream gene is modulated by a factor of ∼100-fold. This is in line with previous studies that report specific cases of measured translational coupling effect of up to 40-fold.[20] In our study we characterized some of the factors affecting the coupling magnitude; specifically we analyzed the gene sequence, position in the operon, and the distance between the stop codon of the upstream gene to the start codon of the downstream gene.

In a previous study, translational coupling was reported to have different characteristics depending on the position of the genes in the operon. Oppenheim and Yanofsky measured translational coupling at various locations in the *trp* operon and reported different translational enhancement for different gene locations.[26] However, as they studied a native operon, they compared between different genes at the different locations, and therefore it remains unclear if the difference in the coupling enhancement is due to the gene location or to the difference in the genes themselves. In our system, we compared between the same genes in the different locations and found that the level of coupling enhancement in our case is unaffected by the location in the operon.

Using a similar synthetic fluorescent reporter gene system, Lim, Lee and Hussein reported that a gene's expression increases with the length of the operon and as its position is moved upstream from the end of the operon.[11] It is difficult for us to observe such an effect in our system since the translational coupling effect is more than an order of magnitude larger than the effect they report, and thus translational coupling dominates the results.

We found that the translation efficiency of one gene affects the translational level of downstream genes also indirectly through a cascading effect. This observation should be kept in mind. For instance, when knocking out a gene from an operon, one might think that only the KO gene is affected, but actually the expression level of all downstream genes in the operon might be reduced as well, potentially causing a strong phenotype by itself. This is akin in effect (but not in mechanism) to the polarity effect,[28,29] known for many years in bacterial genetics where an upstream gene affects transcription rather than translation.

Two main models were previously proposed to explain the coupling phenomenon. The first suggested that translation of the first gene causes a high localized concentration of ribosomes around the translation initiation site of the second gene, thus elevating its translation. Alternatively, it was proposed that a polycistronic message, in the absence of translation, forms an inhibitory secondary structure that sequesters the translation initiation site of the coupled gene; efficient translation of the upstream gene opens up this structure and enables efficient translation of the second gene. Notably, these models are not mutually exclusive. Some observations exist that fit neither of these models and are yet to be explained.[25]

How do our observations relate to these models? We can only speculate on the basis of limited information. The sequence dependence of translational coupling in our system (Figures 3 and 4) is in line with the secondary structure model prediction and does not fit the pure ribosome concentration model prediction. Additionally, preliminary mRNA secondary structure analysis of the intergenic regions of our constructs hints to a possible link between the mRNA secondary structure and the occurrence of coupling. However, the prediction power of the analysis is limited due to the need to use a relatively long stretch of mRNA (for analysis see Supplementary Figure S6). On the other hand, in the early stop codon experiment, we observed a gradual decrease in the fluorescence of the downstream gene as a function of the distance between the stop codon of the upstream gene and the start codon of the downstream gene, with no sudden drops in translational enhancement (Figure 8). This goes beyond the secondary structure model predictions and can be better explained by the ribosome recruitment mechanism, although we do not have any direct measurement for an elevation of ribosome concentration at the start codon of the downstream gene. We interpret this to indicate that the actual modulation is a combination of several mechanisms; specifically, we find that no single mechanism previously proposed can explain all observed results and that we see indications consistent with the two leading mechanisms as playing part in the translational modulation.

Our study supplies the synthetic biology community with an enhanced characterization of translational coupling. The effect can be substantial, modulating gene expression level by over 10-fold, and can cascade across an operon. We think this should be borne in mind when designing synthetic operons and can be used to interpret expression levels in natural and man-made operons.

METHODS

Bacterial Strains and Reagents. The bacterial strain used for cloning and construct assembly was *E. coli* DH5α, and the

dx.doi.org/10.1021/sb400002n | *ACS Synth. Biol.* 2013, 2, 327−336

bacterial strain used for fluorescent measurements was *E. coli* K12 MG1655. All primers were synthesized by Sigma Aldrich, Israel. PCR reactions were performed using Phusion polymerase (Finnzymes, Finland). Restriction enzymes were purchased from New England BioLabs (Beverly, MA, USA) unless stated otherwise.

Choosing a Compact Set of RBS Sequences To Span Expression Space. In order to find a small set of RBS sequences that spans a large range of translation levels we follow the work of Salis et al.,[5] as also described in Zelcbuch et al.[27] Briefly, we computationally calculated the expected translation rate of each of the RBS sequences attached to various genes (Salis Lab: The Ribosome Binding Site Calculator[30]). We chose RBS sequences whose strengths are the least affected from the downstream sequence. From this limited set we picked 6 RBS sequences, denoted RBS^A for strongest RBS to RBS^F for weakest RBS, which span the largest expression space experimentally (for RBS sequences see Figure 1, for RBS calculator results see Supplementary Table S1).

Flanking "Insulator" Sequences. Since the sequences flanking the RBS can affect expression levels, we decided to place so-called insulator sequences, a constant sequence of ∼20 bp located upstream and downstream of each of the RBSs. Such insulator sequences have been previously reported to be effective in reducing the effect of flanking sequences in the case of promoters.[31,32] The upstream insulator sequence was taken to be a 19 bp sequence not natively found in *E. coli*: TAATAGAAATAATTTTGTTTAACTTTA, while the downstream insulator sequence was taken to be a 21 bp sequence coding for a 6His-tag that can also be utilized for a variety of downstream applications: ATGCATCATCACCATCACCAC (Figure 1).

Plasmids. DNA manipulation and assembly processes were conducted on the pRBS backbone plasmids, which contain no designated promoter. Once the assembly process was completed, the resulting product was subcloned into an expression plasmid, pSB4K5:Ptac. This plasmid was derived from pSB4K5, a BioBrick standard vector with low copy pSC101 replication origin (BioPart: BBa_I50042) and kanamycin antibiotic resistance marker (BioPart: BBa_P1003). LacIq Brick (BioPart: BBa_C0012) and a Tac promoter (BioPart: BBa_K301000) were assembled on pSB4K5 upstream to the multiple cloning site using standard assembly methods to yield pSB4K5:Ptac (for a plasmid map see Supplementary Figure S1).

Operon Assembly Process. For the construction of the different RBS modulated synthetic operons we utilized a combinatorial method described by Zelcbuch et al.[27] Briefly, we first subcloned the coding sequence (for genes sequence see Supplementary Table S2) into several linearized pRBS vectors, each vector containing a distinct RBS sequence upstream to the cloning site. We then performed iterative assembly steps, where at each step an additional RBS modulated coding sequence was added along the operon. Once the assembly process was completed, the resulting product was subcloned into an expression plasmid.

Fluorescence Measurements of Reporter Proteins. *E. coli* K12 MG1655 cells were transformed with expression plasmid carrying the desired operon and grown overnight in LB agar plates. Single colonies were then picked and grown overnight at 37 °C in a 96-well plate (Thermo Scientific Nunclon Sterile 96 Well Plates with Lid 167008) containing M9 media supplemented with 0.2% glucose and kanamycin (34 μg/mL). After overnight incubation, cells were diluted (1:2500) and incubated in an automated robotic platform (Evoware II, Tecan). Every 20 min the plate was transferred by a robotic arm into a multiwell fluorimeter (Infinite M200-pro, Tecan). In each measurement OD was measured at 600 nm, RFP was measured by excitation at 587 nm and emission measurement at 620 nm, YFP was measured by excitation at 520 nm and emission measurement at 555 nm, and CFP was measured by excitation at 433 nm and emission measurement at 475 nm.

Could OD_{600} Measurements Be Influenced by the Absorbance of RFP? The absorption spectrum of the RFP fluorescent protein reporter used in this study is centered at 587 nm and is in danger of affecting the optical density at 600 nm, which we used as a proxy for the number of cells (OD_{600}). If this effect is large, the expression level of RFP might significantly influence the OD_{600} measurements. Following this concern, we first checked our experimental measurements specifically for such an influence and did not see any indication of such an effect in our measurements. Furthermore, we performed a back of the envelope calculation of the expected impact of RFP in relation to the cells' OD. We briefly describe the steps of the calculation: We start with the extinction coefficient of the RFP fluorophore (which is equivalent to a cross section for absorption) and perform unit conversions so that we can compare to the cross section for absorption of cells (that lead to the optical density at 600 nm): extinction coefficient RFP $\approx 7 \times 10^4$ M^{-1} cm^{-1} (see, e.g., BNID 106879) $\approx 7 \times 10^4/6 \times 10^{23}$ L $molecule^{-1}$ $cm^{-1} \approx 10^{-19}$ L $molecule^{-1}$ $cm^{-1} \times (1000\ cm^3\ L^{-1}) \approx 10^{-16}$ cm^2 $molecule^{-1}$.

Now we use the number of cells that give an OD of 1 and take into account that an OD of 1 is an extinction coefficient of about 2 cm^{-1}. An OD_{600} of 1 is 10^8–10^9 bacterial cells mL^{-1} (see, e.g., BNID 104831), and we use the lower value as a more stringent test. Therefore, the extinction coefficient of cells, 2 cm^{-1} OD^{-1}, translates to about 10^{-8} cm^2 $cell^{-1}$.

So to have the two effects comparable requires about 10^{-8} cm^2 $cell^{-1}/10^{-16}$ cm^2 $molecule^{-1} = 10^8$ molecule $cell^{-1}$ (i.e., 100 million RFP copies in each cell). Given that there are only about 10^6–10^7 proteins in total in a bacterial cell of volume ∼1 μm^3, we conclude that it is impossible for the RFP expression to have any significant effect on the OD_{600} measurement even for high levels of expression.

For other cells such as yeast or mammalian cells the number of fluorophore molecules per cell can increase but then also the cross section of the cell increases (i.e., number of cells per OD decreases), and so we predict that the expression of such fluorophores will have a negligible effect on OD_{600} for all cell types we can think of.

Data Analysis. Raw data of optical density and fluorescence was background corrected by subtracting wells containing medium with no cells. Because of the large required dynamic range, we could not analyze wells with weak RBS at low bacteria concentrations. We therefore chose to work at mid to late exponential phase. We analyzed cells with an OD_{600} value of ∼0.1 as measured by the plate reader after media subtraction, equivalent to OD_{600} of ∼0.3 for standard 1 cm path length spectrophotometers. For each measurement point, the RBS activity was defined as the increase of fluorescence during a time window of 1 h centered at the measurement's time divided by the average OD measured during that time:

$$A(T) = \frac{F(T+\tau) - F(T-\tau)}{\int_{T-\tau}^{T+\tau} \mathrm{OD}(t)\,\mathrm{d}t}$$

where A is the RBS activity, F is the fluorescence measurement, and τ = 30 min. The result reflects the increase in fluorescence during 1 h divided by the number of cells contributing to this change. Mean activity was calculated by averaging over 5 measurements around OD 0.1 for each sample:

$$\tilde{A} = \frac{\sum_{i=1}^{i=5} A(t_i)}{5}$$

All analysis steps were performed using custom Haskell software.

RT-qPCR Measurements. Total RNA was extracted from *E. coli* K12 MG1655 cells growing in LB media using a High Pure RNA Isolation kit (Roche Applied Science), followed by reverse transcription using Transcriptor Universal cDNA Master Kit (Roche Applied Science). Real-time PCR was performed on a StepOne Real-Time PCR System (Applied Biosystems), using Fast SYBR Green Master Mix (Applied Biosystems). Results were analyzed in the StepOne software v2.1. using the ddCT method with Chloramphenicol gene as the endogenous control. The primers used for analysis: for YFP, AGCACGACTTCTTCAAGTCC and TGTCGGCCATGATATAGACG; for RFP, CCCCGTAATGCAGAAGAAGA and TCTTGGCCTTGTAGGTGGTC; and for Chloramphenicol, AGCACGACTTCTTCAAGTCC and ACCAGCTCACCGTCTTTC. Every construct was measured across five biological repeats, which had three technical repeats each.

■ ASSOCIATED CONTENT

Supporting Information

RBS calculator results for RBS sequences used in the work, fluorescent genes sequence, plasmid map, the two methods for translational coupling analysis, additional results from the C−Y−R library (supplementing the data shown in Figure 6 and 7), and mRNA measurements of the $YFP^{A/C}$-RFP^{C} 850 (no ATG). This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*Tel: +972-8-9344540. Fax: +972-8-9344540. Email: ron.milo@weizmann.ac.il.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors would like to thank Danna-Li Davidovich, Ido Shani, and Orel Rivni for helping with the experimental work; Yehudit Zohar, Avi Falmholz, Oren Yishai, Shira Amram, Arren Bar Even, Leeat Keren, Sivan Navon, Hassan Massalh, Gur Pines, and Alon Wellner for helpful discussions; and Chaim Karp for proofreading. This work was supported by the European Research Council (260392−SYMPAC), Israel Science Foundation (Grant 750/09) and the Helmsley charitable trust. R.M. is the incumbent of the Anna and Maurice Boukstein Career Development Chair in Perpetuity.

■ REFERENCES

(1) Nandagopal, N., and Elowitz, M. B. (2011) Synthetic biology: integrated gene circuits. *Science 333*, 1244−8.
(2) Benenson, Y. (2012) Biomolecular computing systems: principles, progress and potential. *Nat. Rev. Genet. 13*, 455−68.
(3) Canton, B., Labno, A., and Endy, D. (2008) Refinement and standardization of synthetic biological parts and devices. *Nat. Biotechnol. 26*, 787−93.
(4) Endy, D. (2005) Foundations for engineering biology. *Nature 438*, 449−53.
(5) Salis, H. M., Mirsky, E. a, and Voigt, C. a. (2009) Automated design of synthetic ribosome binding sites to control protein expression. *Nat. Biotechnol. 27*, 946−50.
(6) Pfleger, B. F., Pitera, D. J., Smolke, C. D., and Keasling, J. D. (2006) Combinatorial engineering of intergenic regions in operons tunes expression of multiple genes. *Nat. Biotechnol. 24*, 1027−32.
(7) Kudla, G., Murray, A., Tollervey, D., and Plotkin, J. (2009) Coding-sequence determinants of gene expression in Escherichia coli. *Science*, 255−58.
(8) Mutalik, V. K., Guimaraes, J. C., Cambray, G., Lam, C., Christoffersen, M. J., Mai, Q.-A., Tran, A. B., Paull, M., Keasling, J. D., Arkin, A. P., and Endy, D. (2013) Precise and reliable gene expression via standard transcription and translation initiation elements. *Nat. Methods 10*, 354−60.
(9) Zheng, Y., Szustakowski, J. D., Fortnow, L., Roberts, R. J., and Kasif, S. (2002) Computational identification of operons in microbial genomes. *Genome Res.*, 1221−30.
(10) Okuda, S., Kawashima, S., Kobayashi, K., Ogasawara, N., Kanehisa, M., and Goto, S. (2007) Characterization of relationships between transcriptional units and operon structures in Bacillus subtilis and Escherichia coli. *BMC Genomics 8*, 48.
(11) Lim, H., Lee, Y., and Hussein, R. (2011) Fundamental relationship between operon organization and gene expression. *Proc. Natl. Acad. Sci. U.S.A.*, DOI: 10.1073/pnas.1105692108.
(12) Smolke, C. D., and Keasling, J. D. (2002) Effect of gene location, mRNA secondary structures, and RNase sites on expression of two genes in an engineered operon. *Biotechnol. Bioeng. 80*, 762−76.
(13) Løvdok, L., Bentele, K., Vladimirov, N., Müller, A., Pop, F. S., Lebiedz, D., Kollmann, M., and Sourjik, V. (2009) Role of translational coupling in robustness of bacterial chemotaxis pathway. *PLoS Biol. 7*, e1000171.
(14) Schümperli, D., McKenney, K., Sobieski, D. a, and Rosenberg, M. (1982) Translational coupling at an intercistronic boundary of the Escherichia coli galactose operon. *Cell 30*, 865−71.
(15) McCarthy, J., and Gualerzi, C. (1990) Translational control of prokaryotic gene expression. *Trends Genet. 6*, 78−85.
(16) Berkhout, B., and van Duin, J. (1985) mechanism of translational coupling between coat protein and replicase genes of RNA bacteriophage MS2. *Nucleic Acids Res. 13*, 6955−6967.
(17) Van de Guchte, M., Kok, J., and Venema, G. (1991) Distance-dependent translational coupling and interference in Lactococcus lactis. *Mol. Gen. Genet 227*, 65−71.
(18) Praszkier, J., Wilson, I. W., and Pittard, a J. (1992) Mutations affecting translational coupling between the rep genes of an IncB miniplasmid. *J. Bacteriol. 174*, 2376−83.
(19) Bennett, P. M., de la Cruz, F., and Grinsted, J. (1983) Opening the closed ribosome-binding site of the lysis cistron of bacteriophge MS2. *Nature 305*, 741−743.
(20) Rex, G., Surin, B., Besse, G., Schneppe, B., and McCarthy, J. E. (1994) The mechanism of translational coupling in Escherichia coli. *J. Biol. Chem. 269*, 18118−27.
(21) Spanjaard, R. A., and van Duin, J. (1989) Translational coupling in the presence and absence of a Shine and Dalgarno sequence. *Nucleic Acids Res. 17*, 5501−5508.
(22) Schoner, B. E., Belagaje, R. M., and Schoner, R. G. (1986) Translation of a synthetic two-cistron mRNA in Escherichia coli. *Proc. Natl. Acad. Sci. U.S.A. 83*, 8506−10.
(23) Adhin, M. R., and Van Duin, J. (1990) Scanning model for translational reinitiation in eubacteria. *J. Mol. Biol. 213*, 811−8.

(24) Baughman, G., and Nomura, M. (1983) Localization of the target site for translational regulation of the L11 operon and direct evidence for translational coupling in Escherichia coli. *Cell 34*, 979−88.
(25) Little, S., Hyde, S., Campbell, C. J., Lilley, R. J., and Robinson, M. K. (1989) Translational coupling in the threonine operon of Escherichia coli K-12. *J. Bacteriol. 171*, 3518−22.
(26) Oppenheim, D. S., and Yanofsky, C. (1980) Translational coupling during expression of the tryptophan operon of Escherichia Coli. *Genetics 95*, 785−95.
(27) Zelcbuch, L., and Antonovsky, N. (2013) Spanning high-dimensional expression space using ribosome-binding site combinatorics. *Nucleic Acids Res.*, 1−8.
(28) De Smit, M. H., Verlaan, P. W. G., Van Duin, J., and Pleij, C. W. A. (2009) In vivo dynamics of intracistronic transcriptional polarity. *J. Mol. Biol. 385*, 733−47.
(29) Yarchuk, O., Guillerez, J., and Dreyfus, M. (1992) Interdependence of translation, transcription and mRNA degradation in the lacZ gene. *J. Mol. Biol. 226*, 581−96.
(30) https://salis.psu.edu/software/
(31) Davis, J. H., Rubin, A. J., and Sauer, R. T. (2011) Design, construction and characterization of a set of insulated bacterial promoters. *Nucleic Acids Res. 39*, 1131−41.
(32) Lou, C., Stanton, B., Chen, Y.-J., Munsky, B., and Voigt, C. A. (2012) Ribozyme-based insulator parts buffer synthetic circuits from genetic context. *Nat. Biotechnol. 30*, 1137−42.