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