Cell

Quantifying Absolute Protein Synthesis Rates Reveals Principles Underlying Allocation of Cellular Resources

Gene-Wei Li,[1,2,3,*] David Burkhardt,[2,4] Carol Gross,[2,4,5] and Jonathan S. Weissman[1,2,3,*]
[1]Department of Cellular and Molecular Pharmacology, Howard Hughes Medical Institute
[2]California Institute of Quantitative Biosciences
[3]Center for RNA Systems Biology
[4]Department of Microbiology and Immunology
[5]Department of Cell and Tissue Biology
University of California, San Francisco, San Francisco, CA 94158, USA
*Correspondence: gene-wei.li@ucsf.edu (G.-W.L.), weissman@cmp.ucsf.edu (J.S.W.)
http://dx.doi.org/10.1016/j.cell.2014.02.033

SUMMARY

Quantitative views of cellular functions require precise measures of rates of biomolecule production, especially proteins—the direct effectors of biological processes. Here, we present a genome-wide approach, based on ribosome profiling, for measuring absolute protein synthesis rates. The resultant *E. coli* data set transforms our understanding of the extent to which protein synthesis is precisely controlled to optimize function and efficiency. Members of multiprotein complexes are made in precise proportion to their stoichiometry, whereas components of functional modules are produced differentially according to their hierarchical role. Estimates of absolute protein abundance also reveal principles for optimizing design. These include how the level of different types of transcription factors is optimized for rapid response and how a metabolic pathway (methionine biosynthesis) balances production cost with activity requirements. Our studies reveal how general principles, important both for understanding natural systems and for synthesizing new ones, emerge from quantitative analyses of protein synthesis.

INTRODUCTION

Protein biosynthesis is by far the largest consumer of energy during cellular proliferation; translation by ribosomes is estimated to account for ~50% of the energy consumption of a rapidly growing bacterial cell and ~30% of that for a differentiating mammalian cell (Buttgereit and Brand, 1995; Russell and Cook, 1995). The tremendous cost associated with protein synthesis makes it a key step for regulating diverse cellular functions. Therefore, determining how a cell allocates its synthesis capacity for each protein provides foundational information for systems biology.

A fundamental question is whether it is necessary for the cell to exert tight control over the synthesis of individual protein components. For example, the levels of stoichiometric components of protein complexes could be established by differential degradation of excess subunits (Blikstad et al., 1983; Lehnert and Lodish, 1988), rather than by precise synthesis. Moreover, precise control of steady-state protein abundance may not be critical for the performance of cellular circuits. The architectures of several signaling and metabolic pathways have been shown to be robust against variation in protein levels through posttranslational feedback (Alon et al., 1999; Barkai and Shilo, 2007; Batchelor and Goulian, 2003; Hart et al., 2011; Shinar et al., 2007; von Dassow et al., 2000). It remains to be explored whether these posttranslational mechanisms are the dominant strategy for maintaining proper functions or are simply fail-safe mechanisms added on to fine-tuned protein synthesis. More generally, defining such design principles is key to both understanding and manipulating quantitative behavior of a cell.

Efforts to monitor protein synthesis rates at the global level have mainly relied on pulsed metabolic labeling followed by 2D gel electrophoresis or, more recently, by mass spectrometry (Dennis, 1974; Lemaux et al., 1978; Schwanhäusser et al., 2009). Although relative changes in synthesis rates for the same protein are attainable (Selbach et al., 2008), absolute rates are more difficult to evaluate. Additionally, the precision of pulsed metabolic labeling is limited by requirement for nutrient shifts, which affect instantaneous rates of protein synthesis. Alternative methods for expression profiling by determining global mRNA levels (e.g., by high-density microarrays or RNA sequencing [RNA-seq]) do not report the extensive regulation present at the level of translation. These constraints point to a need for a label-free method with unbiased and deep coverage of cellular proteins.

Ribosome profiling—deep sequencing of ribosome-protected mRNA fragments—directly captures protein synthesis in natural settings (Ingolia et al., 2009). It is a general tool for monitoring expression as well as enabling identification of novel translational events (Brandman et al., 2012; Brar et al., 2012; Ingolia et al., 2011; Li et al., 2012; Oh et al., 2011; Stern-Ginossar

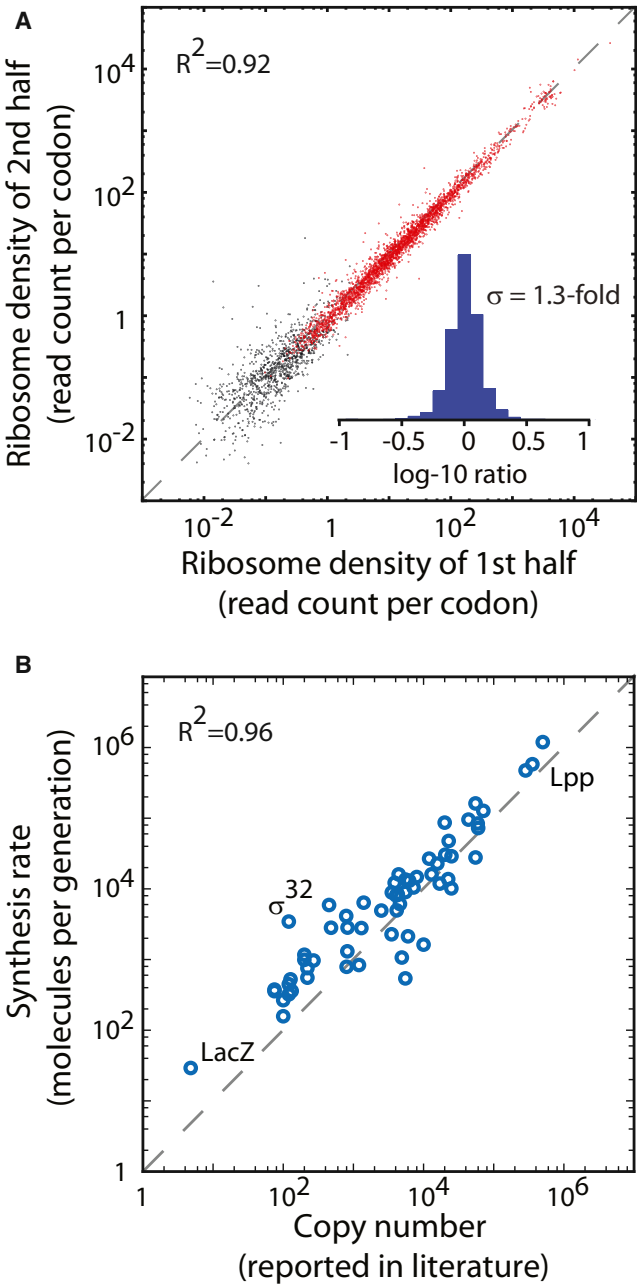

Figure 1. Absolute Quantification of Protein Synthesis Rates

(A) Effect of translational pausing on average ribosome density. Average ribosome density is plotted for the first and second half of each gene. The Pearson correlation for genes with at least 64 reads aligned to both halves (red) is $R^2 = 0.92$. The inset shows the distribution of the fold difference between the second and the first halves (n = 2,870; SD, 1.3-fold).

(B) Agreement between published protein copy numbers and absolute synthesis rates. The copy numbers of 62 proteins that have been individually quantified in the literature are plotted against the absolute protein synthesis rates (Pearson correlation, $R^2 = 0.96$).

See also Figures S1 and S2 and Tables S1 and S2.

et al., 2012). Here, we exploited the ability of ribosome profiling to provide quantitative measurements of absolute protein synthesis rates, covering >96% of cellular proteins synthesized in a single experiment. For stable proteins in bacteria, we then estimated absolute protein copy numbers.

This analysis revealed precise tuning of protein synthesis rates at the level of translation, including a broadly used "proportional synthesis" strategy in which components of multiprotein complexes are synthesized with ratios that quantitatively reflect their subunit stoichiometry. Optimized translation rates are also prevalent among members of functional modules—differential expression pertinent to their functional hierarchy, i.e., when the activity of one member is controlled by the other, was widely observed in our data set. The protein copy numbers inferred from synthesis rates also revealed rules that govern the abundance of transcription factors (TFs) and allowed quantitative characterization for the methionine (Met) biosynthesis pathway, for which we identified a bottleneck enzyme whose expression level is optimized for maximal growth rate. More broadly, our approach and data sets provide a foundation for quantitative understanding of both cellular physiology and precise biological engineering.

RESULTS

Genome-wide Measurement of Absolute Protein Synthesis Rates and Protein Copy Numbers

The ribosome-profiling approach involves freezing of cellular translation followed by digestion of all mRNA regions that are not protected by the ribosome (Ingolia et al., 2009, 2012). Each ribosome-protected mRNA fragment is then identified by massively parallel next-generation sequencing (Ingolia et al., 2009, 2012). Because each ribosome is producing one protein molecule, the rate of protein synthesis is proportional to the ribosome density of a given gene as measured by the footprint density (number of footprint per unit length of the gene), provided that all ribosomes complete a full-length protein and have similar average rates of elongation across genes. Both criteria are broadly met in our data set. During exponential growth in *E. coli*, there is little drop-off in ribosome density for the vast majority of genes (Li et al., 2012; Oh et al., 2011) (Figure 1A). The few genes that display large drop-off could represent novel events of translational regulation (Figure S1A available online). We have previously demonstrated that rare codons are generally translated at similar speed as abundant codons, indicating that differences in codon usage between transcripts do not cause differences in the average rates of elongation (Ingolia et al., 2011; Li et al., 2012). Moreover, sequence-dependent pausing of ribosomes (Li et al., 2012) does not appear to broadly distort the average density of ribosomes along a message because similar ribosome densities are observed in the first and second halves of each gene. Most genes differ by <30% (SD of the mean; Figure 1A). Additionally, correcting for sequence- and position-specific variation in elongation rates has only a modest effect on average ribosome density (Figure S1). Together, these results indicate that local variations in translation speed do not strongly impact synthesis rates measurements based on average ribosome density.