GENE EXPRESSION

Structure versus codon bias

Neil Smith

It is well known that codon bias and gene expression are correlated. The established explanation is that mRNAs with a high codon adaptation index (CAI) — that is, with a high number of 'preferred' codons — are translated more efficiently because there are more tRNAs that match the codons. Now, however, Plotkin and colleagues show that it is mRNA structure, not CAI, that affects expression levels.

The authors generated 154 gene constructs that encoded the same green fluorescent protein (GFP) under the control of a T7 promoter, but in each construct they introduced random synonymous mutations in the third base positions of up to 180 codons. When constructs were put into *Escherichia coli* cells, their fluorescence levels varied 250-fold. However, surprisingly, there was no correlation between expression of the construct and its CAI.

Plotkin and colleagues looked at whether the folding energy of each GFP mRNA correlated with fluorescence. Although the structure of the entire mRNA had no bearing on expression, the folding energy of nucleotide positions –4 to +37 explained over half of the variation: the tighter the folding, the lower the level of expression. These findings support the hypothesis that strong secondary structure at the 5′ end of an mRNA blocks ribosome binding and delays translation initiation.

To test this idea, the authors added an identical stretch of 28 codons with weak mRNA secondary structure to the 5′ end of 72 of the GFP constructs. As expected, the tagged constructs produced consistently high levels of expression. The reduction in translation efficiency for mRNAs with strong folding at the ribosome binding site is consistent with previous studies that suggested that initiation, not elongation, is the rate-limiting step in mRNA translation.

How can these results be reconciled with the well-known link between CAI and expression level? The authors suggest that selection for efficient translation at the global level, rather than at the gene level, has led to an indirect link between expression and CAI among endogenous genes. In their model, high CAI speeds up elongation of a gene but does not affect its expression level. However, faster elongation means that fewer ribosomes are sequestered on the mRNA. This increases the total rate of protein synthesis in the cell, thereby providing a selective advantage in terms of an increased rate of cell growth.

Elizabeth Neame
Copy editor, Nature Reviews Genetics

ORIGINAL RESEARCH PAPER Kudla, G. *et al.* Coding-sequence determinants of gene expression in *Escherichia coli*. *Science* **324**, 255–258 (2009)