Synonymous Mutations Can Affect Protein Folding And Impair Cellular Fitness

March 31, 2020 by Ellen Burns

Synonymous codon substitutions affect the mRNA coding sequence, but the encoded amino acid sequence remains unchanged. Therefore, ostensibly these substitutions do not affect the phenotype and are often ignored in the study of human genetic variation. However, a variety of studies have shown that protein levels, translational accuracy, secretory efficiency, final folding structure and post-translational modifications are regulated by multiple mechanisms.

Synonymous codon action has gradually emerged, and the precise mechanism has yet to be discovered. Studies on the interference of synonymous codon substitution on the co-translational folding mechanism often lack *in vivo* evidence, and usually, rare synonymous codons tend to translate more slowly than ordinary synonymous codons. In addition, rare synonymous codons tend to appear in clusters, many of which are preserved during evolutionary history. The folding rates of many protein secondary and tertiary structures are similar to their synthesis rates, and subtle changes in elongation may also alter the folding mechanism.

Theoretically, synonymous rare codon substitutions reduce translational elongation and can provide more time for the N-terminal portion of the nascent protein to form a stable tertiary structure before the C-terminal portion emerges from the ribosome exit tunnel. Is the extra time good or bad for efficient folding? Cells contain a chaperone network to facilitate protein folding. It is unclear whether altered elongation and co-translational folding mechanisms of synonymous codons interfere with chaperone function.

Recently, Ian M. Walsha and colleagues from the University of Notre Dame published an article in PNAS "Synonymous codon substitutions perturb cotranslational protein folding in vivo and impair cell fitness." They show that synonymous codon changes in the coding sequence of enzymes essential for *E. coli* growth have a significant impact on cell growth. The researchers tested various mechanisms of this growth defect, including changes in folded protein structure, expression levels, enzyme activity, mRNA abundance, and/or cellular stress responses. The findings are compatible with synonymous substitutions that alter the translation elongation pattern, the rationale being that altered co-translational folding mechanisms lead to more degradable structures. These results suggest that synonymous codon alterations can affect protein folding *in vivo* and affect the function of chaperones in protein homeostasis

networks. Thus, the clever use of synonymous codons can have implications for protein design and the interpretation of synonymous mutations associated with the disease.

Most of the current understanding of protein folding mechanisms comes from studies of small proteins that fold reversibly upon dilution from chemical denaturants. However, only a small number of proteins can be stably refolded *in vitro* and remain in a stable state. This suggests that the conformation adopted early in the folding process is essential for successful folding and supports the formation of an early folding intermediate that is different from the conformation formed after dilution from the denaturant. Indeed, there is substantial evidence that chaperones are essential for the successful folding of many proteins *in vivo*. Although it has been hypothesized that synonymous codon changes can alter the extension rate and alter the folding mechanism, to date, no evidence supporting the hypothesis has been found experimentally *in vivo*, possibly due to the auxiliary role provided by chaperones. The results presented here indicate that synonymous codon-induced translational elongation affects folding during the synthesis of nascent CAT polypeptide chains. Although the nascent chain generated using different synonymous codons is still stable and CAT has a trimeric structure, the CAT protein generated by translation using the synonymous Shuf1 mRNA sequence is more susceptible to degradation by the cellular protease ClpXP, which leads to severe cell growth defects. Assuming that the ClpXP ssrA degradation tag is attached to the C-terminus of CAT, most of the degradation may occur only after the release of the translated CAT nascent chain from the nucleosome. Notably, even the native Shuf1-CATssrA protein is more susceptible to degradation than native CATssrA, suggesting that codon-induced perturbations persist for some time after translation and folding are complete. Thus, the buffering effect of the cellular proteostasis network is not sufficient to twist the effect of Shuf1-CAT folding defects on cell growth.

The ssrA tag approach developed here reveals such interferences in other coding sequences, even though these interferences do not lead to eventual changes in protein structure. Recent *in vitro* single-molecule force unfolding experiments have shown that some small, ribosome-bound native folding domains can fold at the ribosome by a similar mechanism. However, as these studies point out, the results measured by molecular tweezers do not capture the transient folding of nascent chains during synthesis, and therefore, in these experiments it is the effect of close proximity of the ribosome surface, rather than co-translational folding of the ribosome surface, that is measured. The folding behavior of a reversible folding model may indeed result in indistinguishable folding behavior during translation. However, most of the model proteins selected for these studies were small, whereas synonymous codon-derived co-translational folds were much larger. The *in vitro* folding mechanism of proteins larger than 175 aa that are retained during co-translational folding is not known. Thus, synonymous codon-derived regulation of elongation rates can play a broad role in efficient folding of larger, and more

complex proteins.

CAT results indicate that synonymous changes in mRNA coding sequences can perturb the folding of protein sequences even in the presence of chaperone molecules, suggesting that mRNA sequences may have evolved with chaperones, thus effectively supporting the folding of the resulting protein structure. Although knowledge of the co-translational folding machinery is still in its infancy, these results suggest that for large or otherwise complex proteins it should be possible to rationally design mRNA coding sequences to improve *in vivo* folding yields and identify disease-associated synonymous codon substitutions that are most likely to adversely affect protein co-translation.