Understanding the Influence of Codon Translation Rates on Cotranslational Protein Folding

Edward P. O’Brien,*,†,‡ Prajwal Ciryam,†,‡ Michele Vendruscolo,† and Christopher M. Dobson*†

†Department of Chemistry, University of Cambridge, Cambridge CB2 1EW, United Kingdom
‡Department of Molecular Biosciences, Northwestern University, Evanston, Illinois 60208, United States

CONSPECTUS: Protein domains can fold into stable tertiary structures while they are synthesized by the ribosome in a process known as cotranslational folding. If a protein does not fold cotranslationally, however, it has the opportunity to do so post-translationally, that is, after the nascent chain has been fully synthesized and released from the ribosome. The rate at which a ribosome adds an amino acid encoded by a particular codon to the elongating nascent chain can vary significantly and is called the codon translation rate. Recent experiments have illustrated the profound impact that codon translation rates can have on the cotranslational folding process and the acquisition of function by nascent proteins. Synonymous codon mutations in an mRNA molecule change the chemical identity of a codon and its translation rate without changing the sequence of the synthesized protein. This change in codon translation rate can, however, cause a nascent protein to malfunction as a result of cotranslational misfolding. In some situations, such dysfunction can have profound implications; for example, it can alter the substrate specificity of an ABC transporter protein, resulting in patients who are nonresponsive to chemotherapy treatment. Thus, codon translation rates are crucial in coordinating protein folding in a cellular environment and can affect downstream cellular processes that depend on the proper functioning of newly synthesized proteins. As the importance of cotranslation translation rates makes clear, a necessary aspect of fully understanding cotranslational folding lies in considering the kinetics of the process in addition to its thermodynamics.

In this Account, we examine the contributions that have been made to elucidating the mechanisms of cotranslational folding by using the theoretical and computational tools of chemical kinetics, molecular simulations, and systems biology. These efforts have extended our ability to understand, model, and predict the influence of codon translation rates on cotranslational protein folding and misfolding. The application of such approaches to this important problem is creating a framework for making quantitative predictions of the impact of synonymous codon substitutions on cotranslational folding that has led to a novel hypothesis regarding the role of fast-translating codons in coordinating cotranslational folding. In addition, it is providing new insights into proteome-wide cotranslational folding behavior and making it possible to identify potential molecular mechanisms by which molecular chaperones can influence such behavior during protein synthesis. As we discuss in this Account, bringing together these theoretical developments with experimental approaches is increasingly helping answer fundamental questions about the nature of nascent protein folding on the ribosome.

INTRODUCTION

The question of how best to understand, model, and predict the influence of codon translation rates on cotranslational folding1−10 (Figure 1a) has recently come to prominence as a result of new experimental strategies3,4,11 and developments in theory and modeling.12−18 We discuss here how physical chemistry, computer simulations, and mathematical modeling can provide useful concepts to describe this phenomenon and how knowledge from these fields can be utilized to model and predict the effect of codon translation rates, which can vary as a result of synonymous codon usage (Figure 1b), on cotranslational folding. For these reasons, it is crucial for the advancement of the "nascent proteome" field (which concerns the proteome-wide behavior of newly synthesized proteins) to bring together the advances being made experimentally with the tools and insights being developed through theoretical and computational investigations.

THE IMPORTANCE OF TIME TO COTRANSLATIONAL PROTEIN FOLDING

A living cell is not in equilibrium with its environment. The foremost implication of this fact is that the behavior of a cell, or a component of it, is determined by the rates of the processes that occur within it rather than by its thermodynamic properties. As it has been put recently, "kinetics can trump thermodynamics" in a system that is out of equilibrium.19 Since the elementary reaction rate of a single-step reaction is the inverse of the average time it takes for this reaction to occur, an equivalent statement is that the properties of a system are determined by the time scales of the processes occurring within that system. Time is therefore a central factor in understanding out-of-equilibrium processes in living cells. For this reason, to understand the impact of changing codon translation rates, it is...
crucial to identify first the essential features of the processes associated with translation and cotranslational folding. The process of translation involves the unidirectional translocation of a ribosome along an mRNA molecule one codon at a time\(^{20}\) (Figure 2). The ribosome decodes the genetic information contained in each codon and covalently attaches the amino acid corresponding to that codon to the growing nascent chain, which moves through a channel that is known as the exit tunnel that is located within the large ribosomal subunit.\(^{21}\) The ribosome exit tunnel is about 10 nm length.\(^{14}\) The biological importance of codon translation rates is also manifested across the genomes of organisms in terms of their codon usage in mRNA molecules. Natural selection has resulted in the presence of particular synonymous codons in particular positions in genes, and in some cases it appears very likely that the result is to coordinate cotranslational folding.\(^{29}\) Experimental support for this concept includes the observation that one of the most successful strategies for designing heterologous protein expression reproduces the codon translation rate profile from the source organism in the recipient organism.\(^{7,8,27,28}\) It is hypothesized that this "codon harmonization" procedure out-of-equilibrium process in a cell. Thus, the underlying reaction rates of codon translation and of domain folding and unfolding have the potential to strongly influence the cotranslational folding behavior of a nascent chain.

We can more clearly define the specific rates that are most relevant to cotranslational folding by considering the details of the process. An elongation cycle of the ribosome, which starts with the P site of the ribosome centered on codon \(i\) and ends with it centered on codon \(i + 1\), involves many intermediate steps, such as tRNA accommodation, peptide bond formation, mRNA translocation, and ejection of the deacylated tRNA from the E site\(^{20}\) (Figure 2). We can subsume these intermediate steps, each with its own elementary reaction rate, into a composite rate \(k_{A,i+1}\), which we call the rate of amino acid addition of the \(i + 1\) residue to the growing nascent chain. This process of amino acid addition extends the nascent chain from \(i\) residues to \(i + 1\) residues in length (Figure 2). The rate \(k_{A,i+1}\) is an important quantity because it defines the average speed at which codon \(i\) is translated.

The other relevant rates are those associated with the process of domain folding at codon \(i\). For example, a domain that folds in a two-state manner and is not influenced by neighboring domains has two distinct rates, \(k_{FU,i}\) and \(k_{FL,i}\), that characterize these intradomain structural transitions at each nascent chain length.\(^{14}\) \(k_{FU,i}\) and \(k_{FL,i}\) are, respectively, the folding and unfolding rates of a domain embedded in a nascent chain that is \(i\) residues long at a given point during its synthesis. These rates depend on the nascent chain length because the chemical environment experienced by the domain changes as the nascent chain is elongated during synthesis.\(^{11,14,23}\) More complex domain folding mechanisms involving intermediates involve an even larger number of rates that can vary at each nascent chain length.\(^{17}\)

This reasoning, based purely on physicochemical principles, suggests therefore that the rates \(k_{A,i+1}\), \(k_{FU,i}\), and \(k_{FL,i}\) have the potential to determine the cotranslational folding behavior of a protein in a cell because of the out-of-equilibrium nature of the translation process.

### EXPERIMENTAL EVIDENCE FOR THE INFLUENCE OF CODON TRANSLATION RATES ON COTRANSLATIONAL FOLDING

There is a substantial amount of evidence that cotranslational folding occurs both in vitro and in vivo.\(^{1,24,25}\) There have also been a smaller but still significant number of experiments probing the consequences for cotranslational folding of changing codon translation rates. Those experiments demonstrate that codon translation rates (i.e., the \(k_{A,i+1}\) values) can determine the probability of cotranslational folding\(^{6}\) and misfolding\(^{4,5,26}\) as well as whether or not a protein ends up functioning properly.\(^{7,8,27,28}\)

The biological importance of codon translation rates is also manifested across the genomes of organisms in terms of their codon usage in mRNA molecules. Natural selection has resulted in the presence of particular synonymous codons in particular positions in genes, and in some cases it appears very likely that the result is to coordinate cotranslational folding.\(^{29}\) Experimental support for this concept includes the observation that one of the most successful strategies for designing synonymous mRNA sequences for heterologous protein expression reproduces the codon translation rate profile from the source organism in the recipient organism.\(^{7}\) It is hypothesized that this "codon harmonization" procedure
maintains the evolutionarily optimized timing of the synthesis of different segments of a protein, thereby maximizing cotranslational folding.

DISTINCTIONS AMONG EQUILIBRIUM, QUASI-EQUILIBRIUM, AND NONEQUILIBRIUM REGIMES OF COTRANSLATIONAL FOLDING

The three rates $k_{i,i+1}$, $k_{i|i}$, and $k_{i|i+1}$ give rise to three broad regimes of cotranslational folding: an equilibrium regime, a quasi-equilibrium regime and a nonequilibrium regime. As discussed below, an out-of-equilibrium process can be either quasi-equilibrium or nonequilibrium in nature (Table 1).

Equilibrium cotranslational folding can be studied in a test tube by stalling indefinitely a ribosome at codon $i$ (i.e., setting $k_{i,i+1} = 0$) and measuring the properties of the nascent chain at that length. Because in this case the nascent chain is not growing, there is no irreversibility in the system, and equilibrium can be achieved. By studying nascent chains arrested at different chain lengths, the equilibrium cotranslational folding curve can be measured. Such equilibrium behavior, where thermodynamics governs the properties of folding at each length, can be used as a reference process against which to compare in vivo cotranslational folding.

Carrying out such comparisons has revealed that two different out-of-equilibrium classifications are possible (Figure 3a). The first is a quasi-equilibrium regime, in which cotranslational folding exhibits equilibrium-like behavior despite being an out-of-equilibrium process because of the irreversible nature of in vivo translation. The other is a nonequilibrium regime, in which cotranslational folding exhibits properties markedly different from those found at equilibrium. A protein can switch between these quasi-equilibrium and nonequilibrium regimes by alteration of the ratio of the folding and codon translation rates, as discussed below. For example, the use of coarse-grained molecular dynamics simulations, in which residues are attached to the growing nascent chain every 60 ms, results in the quasi-equilibrium cotranslational folding of protein G, whereas higher elongation rates result in nonequilibrium cotranslational folding behavior (Figure 3a).

The existence of these three regimes is significant because it can tell us in specific cases whether or not the translation rate is important to cotranslational folding than the thermodynamic properties of the ribosome–nascent-chain complex (RNC). Furthermore, the nonequilibrium regime implies that the cotranslational folding process itself can depend in a sensitive manner on the initial conditions and the history of the RNC, leading to altered folding pathways, structures, and populations from one ribosome molecule to the next in living cells. Such “memory effects” do not occur in the equilibrium and quasi-equilibrium regimes because there is sufficient time for the nascent chain to equilibrate at each nascent chain length and hence “forget” the initial conditions under which it was prepared. A practical implication of this phenomenon is that in a single-molecule experiment involving quasi-equilibrium cotranslational folding, each nascent chain exhibits the same time-averaged behavior at each length $i$ from one ribosome molecule to the next, whereas in the nonequilibrium regime each nascent chain can exhibit different time-averaged behavior.

Table 1. Thermodynamic Classification of Cotranslational Folding Behavior

<table>
<thead>
<tr>
<th>classification</th>
<th>definition</th>
<th>relevance to cotranslational folding</th>
</tr>
</thead>
<tbody>
<tr>
<td>equilibrium</td>
<td>The system under study contains only reversible processes and exhibits properties that do not change over long time periods.</td>
<td>Equilibrium cotranslational folding behavior is exhibited by arrested ribosome molecules that are arrested for times longer than those for all of the other processes.</td>
</tr>
<tr>
<td>quasi-equilibrium</td>
<td>The system consists of one or more irreversible processes that prevent the establishment of equilibrium, yet the system still exhibits equilibrium-like properties.</td>
<td>Quasi-equilibrium cotranslational folding behavior is exhibited when the folding probability at each nascent chain length during continuous translation is comparable to the folding probability on an arrested ribosome.</td>
</tr>
<tr>
<td>nonequilibrium</td>
<td>The system consists of one or more irreversible processes and does not exhibit equilibrium properties. If the properties do not change with time, the system is said to be in a steady (or stationary) state.</td>
<td>Nonequilibrium cotranslational folding behavior is exhibited when the folding probabilities at some nascent chain lengths during continuous translation differ from the folding probabilities on an arrested ribosome.</td>
</tr>
<tr>
<td>out of equilibrium</td>
<td>The system consists of one or more irreversible processes, but whether it exhibits equilibrium-like behavior is not specified.</td>
<td>Knowledge that a protein is being continuously translated in a cell is sufficient information to know that cotranslational folding is out of equilibrium. However, this information is insufficient to determine whether the protein will exhibit quasi-equilibrium or nonequilibrium cotranslational folding behavior.</td>
</tr>
</tbody>
</table>

![Figure 3](image-url)
THEORETICAL TOOLS FOR UNDERSTANDING AND PREDICTING THE INFLUENCE OF CODON TRANSLATION RATES AT THE MOLECULAR AND CELLULAR LEVELS

Significant advances have recently been made in theory and in simulation techniques to understand, model, and predict cotranslational folding behavior at levels ranging from individual proteins to entire proteomes. On the theoretical front, analytical solutions have been determined for chemical reaction schemes representing cotranslational folding that are difficult to solve because of their complexity (Figure 3b). These advances are significant for several reasons. The equations involved enable predictions to be made without having to resort to numerical simulations of the reaction schemes, which can introduce statistical inaccuracies in the calculated quantities. These results also provide a direct means of utilizing the unfolding and folding rates ($k_{\text{FU,i}}$ and $k_{\text{UF,i}}$, respectively) measured on arrested ribosomes to predict the behavior of nascent proteins during continuous translation.  They also allow predictions to be made concerning cotranslational folding scenarios that might occur when codon translation rates are changed. In combination with models for estimating protein folding and unfolding rates, it is possible to use these analytical solutions to predict the cotranslational folding behavior of most of the proteins in the proteome of an organism. Finally, these equations provide a potential means of designing mRNA sequences using synonymous codons to maximize cotranslational folding.

To exemplify these advances, we consider eq 1, which describes the influence of codon translation rates on the ensemble-averaged cotranslational folding probability of a domain that folds via a two-state mechanism:

$$P_k(i) = \sum_{k_{\text{FU,j}}}^{i} \frac{k_{\text{UF,j}}}{k_{\text{FU,j}} + k_{\text{UF,j}}} \prod_{l=1}^{i} \frac{k_{\text{A,j+1}}}{k_{\text{A,j+1}} + k_{\text{UF,i}}}$$

where $i$ is the nascent chain length at a given point during synthesis. To use eq 1, three rates must be known at various nascent chain lengths to make predictions for a specific protein. Although for a 200-residue protein, the requirement could correspond to up to 600 rates, far fewer are needed in practice, as only those rates for chain lengths in the transition region are required. The transition region is defined as the range of nascent chain lengths over which a domain in the protein of interest shifts from a primarily unfolded to a primarily folded conformation at equilibrium (Figure 4). The transition region can only be reached once the domain has emerged from the exit tunnel, and this can be taken into account in eq 1 by setting $k_{\text{UF,i}} = 0$ for chain lengths at which segments of a domain reside inside the tunnel. For a domain that folds in a two-state manner, the size of this transition region can be small, spanning just six residues (Figure 4a). At nascent chain lengths shorter than the onset of this transition region, $k_{\text{FU,i}} \gg k_{\text{UF,i}} \approx 0$ (i.e., $k_{\text{UF,i}}$ is negligible relative to $k_{\text{FU,i}}$), and at lengths greater than this transition region, $k_{\text{FU,i}} \gg k_{\text{UF,i}} \approx 0$ (Figure 4b). Moreover, if correlations among the rates exist in this transition range, even fewer rates need to be measured. In this case, it may be possible to measure, for example, just the rates corresponding to every other chain length in the transition region, estimate the unmeasured rates, and still make accurate predictions using eq 1.

Figure 4. (a) Illustration of the transition region in the cotranslational folding curve (blue region) over which a domain is converted from a predominately unfolded to a predominately folded state at equilibrium. (b) Outside the transition region, one rate (either $k_{\text{FU,i}}$ or $k_{\text{UF,i}}$) predominates and the other is very low, making it unnecessary to measure the low rate in order to make predictions using eq 1. The width of the transition region can vary among proteins, with noncooperative domains that populate intermediates more likely to exhibit larger transition regions. Image reproduced from ref 14.

We note that eq 1 deals only with the elongation phase of translation, not the initiation phase. As the initiation rate determines the quantity of protein synthesized per unit time, eq 1 cannot be used to calculate the translation efficiency of individual transcripts. Furthermore, eq 1 does not account for the rates of mRNA and protein degradation. This fact, however, should not affect the predictions significantly because transcripts being actively translated are less likely to be degraded than those not undergoing translation and cotranslational protein degradation is estimated to be minor, affecting up to 16% of nascent chains in yeast.

On the simulation front, to probe the molecular influence of codon translation rates on the cotranslational folding process, new coarse-grained models of the ribosome have been developed, as conventional all-atom molecular dynamics simulations cannot currently be used to simulate the time scales (tens of seconds) that are relevant to cotranslational folding. As a result of the smoother energy landscape created by coarse graining and the acceleration of dynamics by the low-friction Langevin dynamics approach, coarse graining of the ribosome can effectively model behavior taking place on these time scales.

We discuss here two coarse-grained models that have recently been introduced to simulate the behavior of the ribosome, although several other models on lattices have also been introduced. In both models, ribosomal proteins are represented as either one or two interaction sites per residue, a description based on models previously developed for protein folding in vitro. The two models differ, however, in
their representation of the rRNA. In one model, an all-atom representation of the RNA is used, while in the other model the RNA is represented as up to four interaction sites per nucleotide (Figure 5). Both models treat solvent effects implicitly using Debye–Hückel theory to account for counterion charge screening and Langevin dynamics to account for solvent buffering of the protein and RNA molecules. This type of coarse graining for domains that are relatively small in size (≤100 residues) allows hundreds of cotranslational folding events to be simulated.14,18

![Figure 5](image.png)

**Figure 5.** Coarse-grained model of the *E. coli* ribosome. Ribosomal proteins are represented as one or two interaction sites per residue, and rRNAs are represented as three or four interaction sites. As a consequence, the 50S subunit is reduced from nearly 150 000 atoms to 15 000 interaction sites.18

Such coarse graining has been extended to other cellular components associated with cotranslational folding, including the chaperone Trigger factor that interacts with the nascent chain cotranslationally (Figure 6). In this case, the relevant pairwise binding energies between the components of the ribosome, Trigger factor, and nascent chain were set by determining the nonbonded force-field parameters that reproduced the experimentally measured dissociation constants, $K_D$, between the different components.41,42 We discuss below the insights that these kinetic and simulation models have provided about cotranslational folding.

### THE INFLUENCE OF CODON TRANSLATION RATES ON COTRANSLATIONAL FOLDING CAN BE PREDICTED QUANTITATIVELY

The ability of an analytical solution such as eq 1 to predict accurately the effect of codon translation rates on cotranslational folding on the basis of arrested ribosome data has recently been illustrated.14 Data on $k_{FU,i}$ and $k_{UF,i}$ were generated from explicit coarse-grained molecular simulations of a nascent chain, that of protein G, arrested on an *Escherichia coli* ribosome at different nascent chain lengths (Figure 4b). The values of these rates were then used as the arguments in eq 1 to predict how the folding of this protein would behave during continuous translation; these predictions were then tested against explicit molecular dynamics simulations of continuous translation. The results revealed quantitative agreement between the predictions and the results from the explicit simulations (Figure 7a). Even the effect of changing the translation rate of a single codon on the cotranslational folding curve of this protein was accurately predicted by eq 1 (Figure 7b). The results illustrate in a dramatic manner how measurements11 or estimates15 of $K_{FU,i}$, $K_{UF,i}$, and $k_{A,i+1}$ can be used to predict accurately the influence of codon translation rates on cotranslational folding behavior.

### FAST-TRANSLATING CODONS CAN COORDINATE COTRANSLATIONAL FOLDING BY AVOIDING MISFOLDING

It is commonly assumed that decreasing codon translation rates monotonically increase the probability that a domain folds before the protein is released from the ribosome.1 Indeed, as mentioned previously, there are several experimental reports that support this conclusion. An important question, however, is whether or not this inverse relationship between translation elongation speed and cotranslational folding is a universal phenomenon.

In order to test whether other scenarios are possible, the dynamic regimes that can occur in the analytical models describing the effects of codon translation rates on domains that fold via two-state or three-state cotranslational folding mechanisms involving both on- and off-pathway intermediates were examined.17 This objective was achieved by applying the first-derivative test for monotonicity to the analytical equations.17 The results indicate that, counter to conventional wisdom, slow-translating codons can sometimes decrease
cotranslational domain folding and fast-translating codons can increase it. Within these models, such behavior can occur when nonmonotonic changes in domain stability occur as a function of the nascent chain length or when off-pathway intermediates can be populated during protein synthesis. In the case of domains that fold in an effective two-state manner, speeding up translation through regions of a polypeptide chain where the stability of the folded state of a domain switches from increasing stability to decreasing stability can significantly increase the probability that the domain is folded upon release from the ribosome. For domains that can form an off-pathway intermediate, it was found that speeding up translation through segments of the nascent chain where the intermediate state is stable can increase the final domain folding probability (Figure 8). Thus, for proteins that fold by effective two-state and three-state mechanisms, there are scenarios where fast-translating codons can increase the probability of folding.

The likelihood that intermediates are populated during the folding process in bulk solution increases with increasing protein size, with domains that are less than 80 residues in length typically not populating any intermediate state. In both prokaryotes and eukaryotes, however, protein sizes are on average between 200 and 300 residues, suggesting that the potential to form intermediates could be common. Therefore, nature may have found widespread use for fast-translating codons in coordinating the cotranslational folding of proteins. This hypothesis leads to the prediction that there could be patterns of fast-translating codon usage across the transcriptome of organisms that are consistent with this newly identified function of fast codons. This prediction can be tested by using genomic engineering to systematically search for fast-translating codons for which synonymous substitutions to slower-translating codons significantly decrease the probability of cotranslational folding.
Kinetic modeling of cotranslational folding combined with large biological data sets is opening up new ways to predict cotranslational folding behavior of entire proteomes under varying cellular conditions. This situation is possible at least in part because of the large number of known protein structures and genomic sequences. This information enables reasonable estimates of $k_{A,i+1}$, $k_{FU,i}$, and $k_{UP,i}$ to be made for each protein domain and mRNA molecule in an organism. Using these rates in eq 1 enables the prediction of in vivo cotranslational folding curves for all of the protein domains in a given proteome.

This approach was recently applied to the cytosolic proteome of E. coli to estimate the fraction of the proteome that cotranslationally folds and the degree to which kinetic effects govern cotranslational folding. The results predict that close to 40% of proteins (not weighted by expression levels) in the cytosolic proteome exhibit cotranslational folding at the translation rates found in vivo, with at least one domain in each protein acquiring its stable native structure before the full-length nascent chain is released from the ribosome. In that study, the predicted behavior of individual protein domains was reported, providing a data set against which high-throughput techniques for monitoring cotranslational folding can be compared.

That study also indicated that for the majority of cytosolic proteins, there is likely a deviation between cotranslational folding curves generated on arrested ribosomes ($k_{A,i+1} = 0$), where thermodynamic factors govern the nascent protein behavior, and the curves that occur at in vivo translational rates ($k_{A,i+1} > 0$). For these proteins, cotranslational folding is governed by the translation kinetics at one or more nascent chain lengths. This analysis suggests that around 20% of cytosolic proteins contain domains that fold cotranslationally on an arrested ribosome but fold post-translationally at natural translation rates. That is, synthesis in vivo can be so fast that some domains may not fold cotranslationally despite a thermodynamic driving force to do so. Thus, the translation kinetics can have profound effects on the likelihood of cotranslational folding in vivo.

The basis for the predominance of kinetics in E. coli cotranslational folding is that the time scale of the folding of the majority of domains is longer than the time scale of amino acid addition ($\tau_F > \tau_L$). As a consequence, these domains cannot equilibrate at every nascent chain length and thus are unable to sample effectively the extent of conformational space that is in principle available to them. According to the study in ref 15, larger domains rich in $\beta$-sheet structure tend to exhibit the largest deviations from equilibrium, while small domains containing primarily $\alpha$-helical structure are more likely to exhibit quasi-equilibrium behavior. The reason for this difference is that large $\beta$-sheet domains tend to have much larger $\tau_F$ values than small $\alpha$-helical domains, and this time scale, relative to $\tau_L$, determines the cotranslational folding properties (Figure 9).

The importance of translation kinetics for cotranslational folding has other possible consequences. Some proteins may get kinetically trapped in metastable states during synthesis, which in extreme cases could result in substantially different final structures and assembly pathways from those observed in vitro. Thus, the nonequilibrium nature of cotranslational folding in vivo is likely to lead to richer self-assembly behavior than found in vitro. This scenario warrants further experimental investigation, especially with respect to the probability of nascent chain misfolding and the biological and cellular consequences for the function of newly synthesized proteins. Indeed, it has recently been shown that changing the translation rates of a designed protein in vivo can alter the conformation the protein adopts, lending support to the theoretical result that at in vivo translation rates many proteins can fold under a nonequilibrium regime.

**CONCLUDING REMARKS**

A variety of experiments have established that the rate at which individual codons are translated can strongly influence whether some proteins fold to functional states or misfold and malfunction. As misfolded species often arise from partially folded intermediates, some proteins are particularly sensitive to conformational disruption due to inefficient cotranslational folding. Therefore, determining the points during translation where proteins are likely to remain unfolded or misfold could shed light on the molecular mechanisms of some forms of cellular stress. In this Account, we have summarized recent efforts to develop a molecular perspective and a theoretical framework to understand, model, and predict the influences of codon translation rates on these cellular processes. The approaches being developed and utilized range from techniques in the areas of molecular simulations to kinetic analyses and systems biology.

Each approach makes its own assumptions, which define the range of questions that it can effectively address and form the basis for new developments. Thus, for example, conventional chemical kinetic models are limited in their ability to predict single-molecule behavior, at least in part because of the effects of stochasticity, but hybrid approaches that blend chemical kinetics and a discrete time formalism suggest one way forward. Estimating codon translation rates under varying cellular conditions is also a significant challenge, but ribosome profiling data may be very useful, as the number of “reads”
the signal) from such experiments is a function of the codon
elongation rates.51 Other methods and models of translation
rates may also prove to be very useful.4 It is important to test by
experiment the predictions from these approaches, and we
hope that this review will help to promote quantitative
experimental measurement of the coupling of cotranslational
folding to elongation rates. For example, to our knowledge only
two in vivo cotranslational folding curves have been
experimentally measured.24 Advances in single-molecule
techniques and in vivo imaging methods hold out the promise
that more examples will be reported in the future.

Thus, theory and computation are making vital contributions
to our understanding of cotranslational folding and proteome
metastability in living cells. With the continuing advance of
quantitative, real-time measurements of processes in living cells,
the contributions and necessity of these theoretical approaches
will undoubtedly grow in the future.

ACKNOWLEDGMENTS

We thank John Christodoulou and Lisa Cabrita for very useful
discussions concerning cotranslational folding and Attila Szabo
for stimulating suggestions for modeling approaches. We thank
the Engineering and Physical Sciences Research Council (UK)
for financial support (E.P.O., M.V., C.M.D.) and the National
Science Foundation (USA) for the award of a postdoctoral
fellowsip to E.P.O.

REFERENCES

(1) Komar, A. A. A pause for thought along the co-translational
(2) Cabrita, L. D.; Dobson, C. M.; Christodoulou, J. Protein folding
(3) Zhang, G.; Huhalewska, M.; Ignatova, Z. Transient ribosomal
attenuation coordinates protein synthesis and co-translational folding.
(4) Spencer, P. S.; Siller, E.; Anderson, J. F.; Barral, J. M. Silent
substitutions predictably alter translation elongation rates and protein
(5) Siller, E.; DeZwaan, D. C.; Anderson, J. F.; Freeman, B. C.;
Barral, J. M. Slowing Bacterial Translation Speed Enhances Eukaryotic
(6) Chiti, F.; Dobson, C. M. Protein misfolding, functional amyloid,
(7) Kimchi-Sarfaty, C.; Oh, J. M.; Kim, I. W.; Sauna, Z. E.; Calcagno,
A. M.; Ambudkar, S. V.; Gottesman, M. M. A "silent" polymorphism in
the MDR1 gene changes substrate specificity. Science 2007, 315, 525–
528.
(8) Edwards, N. C.; Jing, Z. A.; Perry, A.; Blaisdell, A.; Kopelman, D.
B.; Fatheke, R.; Plum, W.; Newell, J.; Allen, C. E.; S.; G.; Shapiro, A.;
Okunji, C.; Kost, I.; Schmorn, N.; Grigoryan, V.; Pretyycka, T. M.;
Sauna, Z. E.; Salari, R.; Mandel-Gutfreund, Y.; Komar, A. A.; Kimchi-
Sarfaty, C. Characterization of Coding Synonymous and Non-
Synonymous Variants in ADAMTS13 Using Ex Vivo and In Silico
Sachs, M. S.; Liu, Y. Non-optimal codon usage affects expression,
structure and function of clock protein FRQ. Nature 2013, 495, 111–
115.
(10) Tsai, C. J.; Sauna, Z. E.; Kimchi-Sarfaty, C.; Ambudkar, S. V.;
Gottesman, M. M.; Nussinov, R. Synonymous Mutations and Ribosome
Stalling Can Lead to Altered Folding Pathways and Distinct
(11) Kaiser, C. M.; Goldman, D. H.; Chodera, J. D.; Tinoco, I., Jr.;
Bustamante, C. The ribosome modulates nascent protein folding.
Science 2011, 334, 1723-1727.
(12) Elcock, A. H. Molecular simulations of cotranslational protein
folding: Fragment stabilities, folding cooperativity, and trapping in the
(13) O'Brien, E. P.; Christodoulou, J.; Vendruscolo, M.; Dobson, C.
M. New Scenarios of Protein Folding Can Occur on the Ribosome. J.
(14) O'Brien, E. P.; Vendruscolo, M.; Dobson, C. M. Prediction of
variable translation rate effects on cotranslational protein folding. Nat.
(15) Ciryan, P.; Morimoto, R. I.; Vendruscolo, M.; Dobson, C. M.;
O'Brien, E. P. In vivo translation rates can substantially delay the
(16) O'Brien, E. P.; Hsu, S. T. D.; Christodoulou, J.; Vendruscolo,
M.; Dobson, C. M. Transient Tertiary Structure Formation within the
(17) O'Brien, E. P.; Vendruscolo, M.; Dobson, C. M. Kinetic
modeling indicates that fast-translating codons can coordinate
cotranslational protein folding by avoiding misfolded intermediates.
(18) O'Brien, E. P.; Christodoulou, J.; Vendruscolo, M.; Dobson, C.
M. Trigger Factor Slows Co-translational Folding through Kinetic
Trapping While Sterically Protecting the Nascent Chain from Aberrant
(19) Chandler, D. Self-assembly where kinetics trumps thermody-


