Biochemistry
Viewpoint
Cite This: Biochemistry 2018, 57, 3537−3539
pubs.acs.org/biochemistry

# Messenger RNA Structure Regulates Translation Initiation: A Mechanism Exploited from Bacteria to Humans

Anthony M. Mustoe,*,† Meredith Corley,‡ Alain Laederach,‡ and Kevin M. Weeks*,†

†Department of Chemistry and ‡Department of Biology, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599, United States



Translation initiation is a major rate-limiting step in protein synthesis and is highly regulated in all cells. Messenger RNAs (mRNAs) play a central role in directing this regulation. Through instructions that remain largely cryptic, mRNA transcripts modulate the initiation process to achieve a specific translation efficiency (TE), the amount of protein made from a given mRNA transcript. TE is precisely tuned, can vary significantly depending on cell type, varies by orders of magnitude across different transcripts, and thus constitutes an essential variable in gene expression. Understanding how mRNAs encode their own unique TEs is therefore a fundamental challenge in biology.

Classic studies have shown that mRNAs can encode TE by folding into structures that facilitate or impede translation initiation.[1] The mRNA-binding cleft of the ribosome can accommodate only single-stranded mRNA. Thus, translation initiation requires unfolding of any mRNA structures that overlap the start codon, imposing a structure-dependent energetic penalty on initiation (Figure 1). Synthetic biologists have harnessed these principles to tune the TEs of designed mRNAs over a large dynamic range.

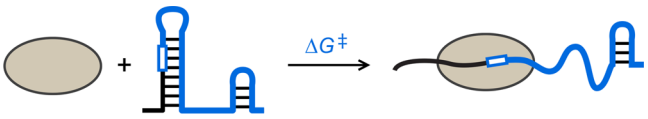


**Figure 1.** General mechanism by which RNA structure regulates translation initiation. The mRNA start codon (empty box) and surrounding sequence must be single-stranded to be accommodated into the mRNA cleft of the ribosomal preinitiation complex (brown). Unfolding the mRNA structure imposes an energetic penalty ($\Delta G^{\ddagger}$) on translation initiation.

Despite these classic studies, the overall importance of mRNA structure in regulating TEs of endogenous genes has remained surprisingly unclear. Until recently, it has been all but impossible to model mRNA structure with high accuracy. Studies of endogenous genes have largely relied on *in silico* structure predictions that have supported only a weak relationship between mRNA structure and TE. Whether this weak relationship reflected biological reality or the shortcomings of RNA structure modeling was a persistent unanswered question. In two recent studies,[2,3] our laboratories used the SHAPE-MaP RNA chemical probing strategy to determine high-confidence, experimentally supported structure models for hundreds of mRNAs, providing a unique opportunity to revisit this question.

In one of these studies,[2] we used a data set of approximately 200 SHAPE-MaP-determined mRNA structures to investigate translation regulation in the simple prokaryote *Escherichia coli*. We initially assumed that it would be straightforward to quantify the influence of structure on TE, but properly addressing this question necessitated a broad evaluation of the mechanism of translation initiation. The ribosomal preinitiation complex recognizes mRNAs via a two-step mechanism. First, the preinitiation complex nonspecifically binds the mRNA via a loosely defined "standby site"; second, the mRNA is unfolded and is accommodated into the mRNA-binding cleft of the ribosomal small subunit, allowing recognition of the Shine-Dalgarno sequence and the start codon (Figure 1).[4] Important details of this second accommodation step were unclear. Specifically, mRNA accommodation might be an equilibrium process wherein the mRNA has time to refold into other low-energy structures, in which case TE should depend on the *equilibrium* free energy of mRNA unfolding ($\Delta G$). Alternatively, accommodation could be a non-equilibrium process during which the mRNA does not have time to refold. In this case, TE will depend on the *non-equilibrium* energy of mRNA unfolding ($\Delta G^{\ddagger}$) (Figure 2A). Our structural data allowed us to estimate $\Delta G$ and $\Delta G^{\ddagger}$ for each gene, which we correlated with TE measurements published by J. Weissman's lab. Strikingly, our data indicated that TE strongly depends on start codon $\Delta G^{\ddagger}$ (non-equilibrium) but not $\Delta G$ (equilibrium). We subsequently validated this conclusion using reporter assays for 29 genes.[2]

These data clearly indicated that mRNA structure tunes gene TE in *E. coli*. Moreover, the finding that TE depends on non-equilibrium mRNA unfolding has broad implications for our understanding of bacterial translation. First, as has been previously theorized, our finding implies that translation initiation is governed by a kinetic competition between mRNA unfolding and dissociation of the ribosome preinitiation complex. Free ribosome complexes are scarce. Because any individual mRNA comprises a small fraction of the cellular mRNA pool, a dissociated ribosome will most likely bind and initiate on another mRNA (Figure 2A).[4] Second, a dependence of TE on RNA unfolding kinetics provides new insight into the function of RNA regulatory motifs such as riboswitches and thermosensors, which regulate translation in response to small molecule binding or changes in temperature. Kinetic competition provides a consistent explanation for how these motifs modulate ribosome initiation, and thus TE, despite only modest observed changes in equilibrium stability.

In a second study,[3] we examined translation of the human gene *SERPINA1*, which encodes the $\alpha$-1-antitrypsin protein. Deficiency of $\alpha$-1-antitrypsin is clinically linked to lung, liver, and inflammatory diseases. *SERPINA1* is a remarkably complex gene that has 11 different splicing isoforms, all of which change

**Received:** April 5, 2018
**Published:** June 12, 2018




DOI: 10.1021/acs.biochem.8b00395
Biochemistry 2018, 57, 3537−3539

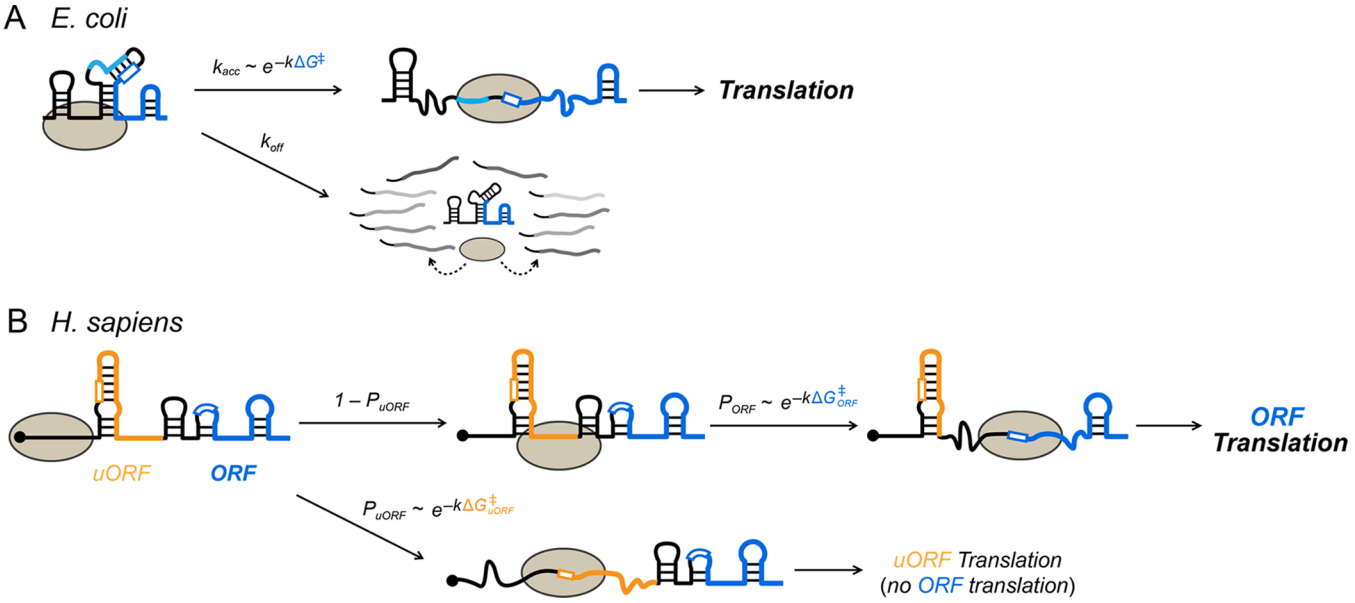


**Figure 2.** Non-equilibrium mRNA unfolding regulates translation in *E. coli* and humans. (A) In *E. coli*, the ribosomal preinitiation complex (brown) first binds the mRNA nonspecifically. Initiation is subsequently governed by mRNA unfolding and accommodation into the mRNA cleft of the preinitiation complex, depending on the non-equilibrium free energy of unfolding the mRNA structure ($\Delta G^{\ddagger}$). The dependence on non-equilibrium mRNA unfolding implies that initiation is governed by a kinetic competition between accommodation of the mRNA into the mRNA-binding site of the ribosome ($k_{acc}$) and ribosome dissociation ($k_{off}$); once dissociated, the preinitiation complex will rapidly bind another mRNA. The mRNA start codon and Shine-Dalgarno sequence are indicated by an empty box and light blue line, respectively. (B) In humans, translation initiation typically proceeds via the scanning mechanism. The preinitiation complex is recruited to the mRNA 5′-cap and subsequently scans along the 5′-untranslated region (5′-UTR) for a start codon. If the 5′-UTR contains a uORF (orange), the preinitiation complex must leak past the uORF to initiate at the primary open reading frame (blue). Otherwise, translation will prematurely initiate and terminate at the uORF. The initiation probability at each start codon ($P_{uORF}$ and $P_{ORF}$) is regulated by the non-equilibrium free energy of unfolding the start codon structure (note that this nomenclature differs slightly from that used in ref 3).

the sequence of the 5′-untranslated region (5′-UTR) but not the protein coding sequence. Each isoform contains different combinations of up to three upstream open reading frames (uORFs), which reduce gene TE by causing premature initiation during the scanning process of eukaryotic translation initiation. Under this model, only ribosome complexes that manage to "leak" past uORFs successfully initiate translation at the primary ORF (Figure 2B). While uORFs are found upstream of approximately 50% of human genes, much about their function remains unknown. Interestingly, existing models of leaky scanning inadequately explained the TE variation of different *SERPINA1* isoforms, prompting us to explore whether 5′-UTR structure affects uORF function. We used SHAPE-MaP experiments to model the structures of each *SERPINA1* isoform and, using the exact same non-equilibrium model that we worked out in *E. coli*, estimated the free energy ($\Delta G^{\ddagger}$, annotated simply as $\Delta G$ in the original paper) required to unfold structure around the start codon of each uORF and the primary ORF. Remarkably, consideration of the start codon $\Delta G^{\ddagger}$ dramatically improved the predictive power of the leaky scanning model for the eukaryotic *SERPINA1* isoforms (Figure 2B). We subsequently validated this conclusion by selectively disrupting structures and observing corresponding changes in TE.[3]

Our analysis of *SERPINA1* strongly implies that mRNA structure regulates start codon recognition of both the primary ORF and any uORFs during eukaryotic scanning via a mechanism similar to that uncovered in prokaryotes. This finding is surprising on two levels. First, the canonical model of eukaryotic scanning posits that 5′-UTR structures are requisitely unfolded as the mRNA is threaded through the mRNA-binding cleft of the ribosome preinitiation complex.[1] However, if all mRNA structure is unfolded, it is unclear how RNA structure could tune start codon recognition, as we have observed and mechanistically validated. Thus, for *SERPINA1*, and potentially many other mRNAs, our findings suggest that this canonical model needs to be revised. Very stable stem loops in a 5′-UTR are known to stall the scanning process. Nevertheless, we speculate that, in specific contexts, stable RNA structures and 5′-UTR sequences occluded by antisense oligonucleotides[5] might be bypassed (or scanned over) without unfolding during the scanning process. Second, the mechanism of translation initiation and regulation is significantly more complex in eukaryotes than in prokaryotes. It is therefore striking that both domains of life appear to exploit the simple regulatory strategy of using mRNA structure to tune TE.

We emphasize that high-quality RNA structural data are an absolute prerequisite for understanding the functions of mRNA structure. In both studies, we repeated our analyses using "no-data" (*in silico* only) structure predictions and observed inconclusive relationships between mRNA structure and TE. Native mRNAs have highly diverse and complex structures that are very challenging to predict. These structures may not be well-conserved or uniquely stable compared to expectations of "random" RNA. Nonetheless, idiosyncratic RNA structures that overlap translation start sites play fundamental roles in tuning the TE of each mRNA.

Our analyses are, of course, not perfect. SHAPE-MaP-guided calculations of RNA folding energies include several approximations.[2] In addition, mRNA unfolding is likely manipulated by physical interactions with the ribosome and other proteins, whereas our studies approximate mRNA unfolding as if it occurs in isolation. Our studies thus far have assumed that mRNAs fold



DOI: 10.1021/acs.biochem.8b00395
*Biochemistry* 2018, 57, 3537−3539

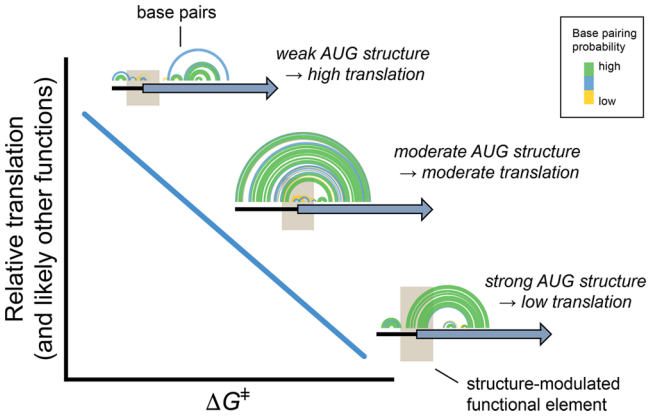


**Figure 3.** mRNAs tune translation via idiosyncratic structures with varied unfolding energies. The brown box indicates the gene start (AUG) region, and arcs illustrate RNA base pairs. Note that this mechanism likely applies to other RNA-mediated processes in which accessibility to a binding partner is tuned by preexisting RNA structure.

into a single structure, whereas many mRNAs likely fold into an ensemble of structures. Tackling such complexities will yield further insight into the mechanisms through which mRNA structure influences translation initiation.

In summary, our studies paint an emerging portrait of start codon structural accessibility as a critical and pervasive regulator of translation initiation. This mechanism is simple, does not require mRNA structure to be particularly well-defined or well-conserved, and notably modulates expression of genes in both prokaryotes and eukaryotes. More broadly, we expect that this general mechanism extends to other RNA-mediated biological processes, with seemingly unremarkable RNA structures likely playing broad roles in tuning interactions between RNA and diverse ligands (Figure 3).

## AUTHOR INFORMATION


### Corresponding Authors

*E-mail: amustoe@unc.edu.

*E-mail: weeks@unc.edu.

### ORCID

Anthony M. Mustoe: 0000-0001-9346-1559

Kevin M. Weeks: 0000-0002-6748-9985



### Funding

A.M.M. is supported by an Arnold O. Beckman postdoctoral fellowship. A.L. is supported by Grants R01 HL111527 and R01 GM101237. K.M.W. is supported by Grants R01 AI068462 (RNA structure methods development) and R35 GM122532 (RNA function discovery).


### Notes

The authors declare the following competing financial interest(s): K.M.W. is an advisor to and holds equity in Ribometrix, to which SHAPE-MaP technologies have been licensed.

## ACKNOWLEDGMENTS


We thank J. Weissman and his lab for publicly sharing their data, and acknowledge the many authors that we could not cite due to space limitations. The authors thank Chase Weidmann for helpful discussions.


## REFERENCES


(1) Kozak, M. (2005) Regulation of translation via mRNA structure in prokaryotes and eukaryotes. *Gene 361*, 13−37.

(2) Mustoe, A. M., Busan, S., Rice, G. M., Hajdin, C. E., Peterson, B. K., Ruda, V. M., Kubica, N., Nutiu, R., Baryza, J. L., and Weeks, K. M. (2018) Pervasive regulatory functions of mRNA structure revealed by high-resolution SHAPE probing. *Cell 173*, 181−195.

(3) Corley, M., Solem, A., Phillips, G., Lackey, L., Ziehr, B., Vincent, H. A., Mustoe, A. M., Ramos, S. B. V., Weeks, K. M., Moorman, N. J., and Laederach, A. (2017) An RNA structure-mediated, posttranscriptional model of human α-1-antitrypsin expression. *Proc. Natl. Acad. Sci. U. S. A. 114*, E10244−E10253.

(4) de Smit, M. H., and van Duin, J. (2003) Translational standby sites: how ribosomes may deal with the rapid folding kinetics of mRNA. *J. Mol. Biol. 331*, 737−743.

(5) Liang, X.-H., Shen, W., Sun, H., Migawa, M. T., Vickers, T. A., and Crooke, S. T. (2016) Translation efficiency of mRNAs is increased by antisense oligonucleotides targeting upstream open reading frames. *Nat. Biotechnol. 34*, 875−880.




DOI: 10.1021/acs.biochem.8b00395
*Biochemistry* 2018, 57, 3537−3539