*Borneo International Journal of Biotechnology (BIJB)*
*Vol. 2 (December 2022), 54 – 64*

*e-ISSN 2716-697X*

# RNA codons govern the mechanism of protein folding through the shape memory effect

**Hou Pin Yoong**

Biomechatronic Laboratory, Faculty of Engineering, Universiti Malaysia Sabah,
Jalan UMS, 88400 Kota Kinabalu, Sabah, Malaysia
Author's email: yoongpin@ums.edu.my
https://doi.org/10.51200/bijb.v2i.2955
**Received: 12 April 2021 | Revised: 24 February 2022 | Accepted: 28 June 2022| Published: 1 December 2022**

## ABSTRACT

The complex protein folding mechanism had been researched during the past half-century, given its potential to offer cures for illnesses caused by viruses and protein misfolding. However, to date, the work remains inadequately successful and mastered, provoking the question of whether researchers are looking at the wrong place for the answer. Specifically, can RNA codons define the protein folding mechanism? This review will first present existing mechanisms for protein folding and their limitations. Then, the logic and evidence supporting the use of a protein folding mechanism governed by RNA codons will be presented. This paper explains protein folding as a shape-memory phenomenon wherein the protein chain memorises the native folded structure. Under the right chemical environment, the protein chain will fold back into its native memorised structure. The RNA codon is the imprint for the natively folded protein shape memory, responsible for programming the native folded structure shape memory onto the protein chain.

**Keywords:** protein folding, shape memory effect, RNA codons, native structure

## INTRODUCTION

The complete knowledge of protein folding mechanisms can provide the complete physical structure of a virus or cancer cell and supports the ability to formulate cures for diseases caused by viruses and protein misfolding (Díaz-Villanueva et al., 2015). Although there are big differences between virus and cancer cells' structure, in general, the translation mechanism is similar in that ribosomes are responsible for synthesising

the polypeptide via RNA translation and later fold to form active protein. These active proteins are parts of the virus or cell such as the membrane that encapsulates the virus or cancer cells. During the protein folding into the membrane structure, receptors are formed on the membrane of the virus or cancer cell. By knowing the exact physical structure and the receptor on the virus or cancer cell, a specific protein can be engineered using the knowledge of protein folding mechanisms to match the receptor on the virus or cancer cell for the means to deactivate it or make it compatible with the T-cell receptor. This will allow the T-cell to recognize the virus or cancer cells and eliminate them. As shown in Figure 1, (a) the T-cell receptors are not compatible with the virus or cancer cell therefore, the virus or cancer cell cannot be recognized nor eliminated by the T-cell. With the complete knowledge of protein folding, a specific protein can be engineered to match the virus or cancer cell receptor and bind to them, this is to either deactivate the virus or a receptor matching the T-cell receptor can be engineered on the other side of the protein for the T-cell to bind on it. The binding of the virus or cancer cell to the T-cell allows the T-cell to recognise and destroy the virus or cancer cell as shown in Figure 1 (b).

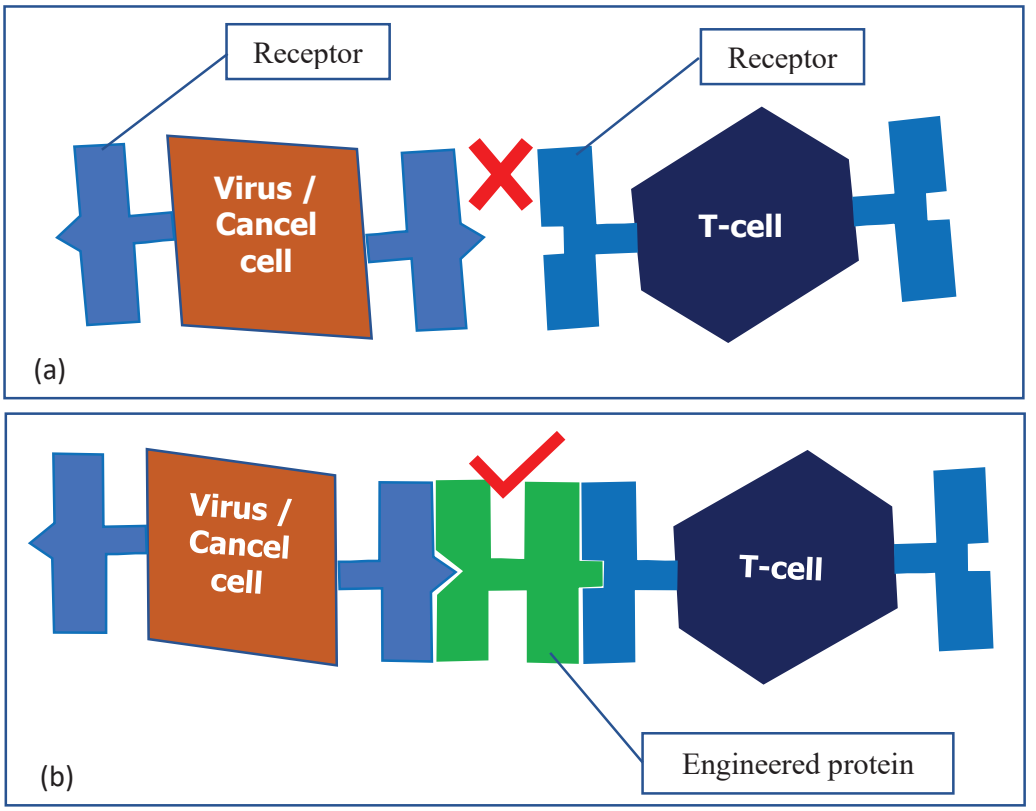


**Figure 1** (a) The T-cell has receptors that do not match with the virus or cancer cell. (b) The engineered protein bridges the virus or cancer cell receptor to the T-cell receptor.

For the virus, the process is more direct since the receptor protein is much more different from the human cells' receptor, which makes it easier to be differentiated. On the other hand, the cancer cell might not be able to be detected due to the mutation might not happen in the part of the codon sequence that forms the cell plasma membrane. If this is the case, then the cancer cell plasma membrane will be the same as other cells and can not be detected and destroyed by the immune system. However, sometimes a small mutation at the codon sequence that forms the plasma membrane causes the cancer cell plasma membrane to be different. But the difference is small and becomes camouflaged to our immune system. Thus, the cancer cell will not be destroyed by the immune system. By knowing the mutation of the codon sequence responsible for producing the plasma membrane the small difference in the cancer cell plasma membrane can be identified and specific proteins can be engineered to attach to it and allow our immune system to identify and destroy it. Thus the understanding of the protein folding mechanism is not limited to either virus or cancer cells. Besides, this article is focused on the protein folding mechanism which is after the maturation process of the RNA.

The knowledge of protein folding not only helps to avoid future world pandemics such as COVID-19, Ebola, Severe Acute Respiratory Syndrome (SARS), etc., but it can also save more lives than we can imagine. For over half a century, researchers have been working to understand the mechanism of protein folding. Yet, all of this hard work returned only limited achievements (Jerath et al., 2016). What has so significantly complicated the path to the understanding of the protein-folding mechanism? This paper will present a brief pass excursion on finding the protein-folding mechanism, and the setback encountered, followed by the proposal and assessment of a new hypothesis suggesting that the RNA codon governs the protein-folding mechanism through the shape memory effect. A conclusion and possible future works are also presented in this paper.

## ANFINSEN'S DOGMA

In 1961, Anfinsen et al. proposed that the native structure of the folded protein achieved the lowest Gibbs free energy level based on the interatomic interactions of the amino acid sequence in the protein chain (Anfinsen et al., 1961). Anfinsen claimed the natural protein chain folded formation known as the native structure always achieves the lowest degree of energy formation (Anfinsen, 1973). Initially, this hypothesis seems to explain the ability of the protein chain to fold back into a native structure after unfolding. However, other research (Levinthal, 1969) has observed that certain native structures consist of folding arrangements that do not comply with the Anfinsen lowest-energy hypothesis. Generally, the native structure can exist simply by achieving a metastable state that is sustainable in the natural biological environment.

Meanwhile, the noncomplier to the lowest-energy structure can exist in many configurations, resulting in an exceptionally large number of possible combinations for the protein chain. Overcoming this, researchers have developed programs to predict the final structure by elucidating the most likely folding pattern. Some longer-chain proteins may require thousands of years of computation time to discern how the native protein folds. Various selection rules had been applied to shorten the guessing time. However, these methods have yet to yield a satisfying result (Onuchic & Wolynes, 2004; Dobson, 2004; Baker, 2000). Based on the discovery of the protein native structure existing without complying with the lowest energy form, the further exploration of Anfinsen's hypothesis became impossible due to the exceptionally large probability of the protein folding orientation as addressed by Levinthal's work and later known as Levinthal's paradox.

## LEVINTHAL'S PARADOX

Levinthal's work has shown that the exceptionally large possibility of protein folding led to an enormously lengthy mean first-passage time, which is the time needed to possibly achieve the native folding structure. Zwanzig et al. (1992) considered that a penalty function to show the mean first-passage time could be reduced to seconds. This prior work might suggest the realisable time for obtaining the optimal possible folding combination for the protein chain. However, a few questions persist. First, is the protein folding mechanism in nature governed by the same penalty function as proposed by Zwanzig et al.? Second, what is the probability of achieving successful folding to the native structure? Based on Levinthal's work, the probability of a complete protein chain folding exactly to the native form is about 1 from $10^{300}$, even for a protein chain that only consists of 100 protein molecules (Finkelstein et al., 2017). In other words, nearly 100% of the protein chain exhibits a risk of misfolding, however, this situation is in contradiction with the existing biological system as reality suggests the possibility of diseases caused by the protein-chain misfolding is low. For example, Alzheimer's disease (Chiti, 2017; Dobson, 2003) was reported to affect only 0.63% of the world's population as of 2015 (Vos, 2016). This evidence strongly supports that protein folding based on random formations is inconsistent with reality.

# COMBINATION OF PROTEIN TOPOLOGICAL AND ENERGY FUNCTION FOR PROTEIN NATIVE-STRUCTURE PREDICTION

Levinthal's paradox adopts a very important view that protein-chain folding cannot happen randomly (Tompa, 2011). Based on this understanding, researchers began looking into specific folding patterns and used them as possible predictors for disease (Karplus, 1997). Some researchers simplified the protein model into a two-dimensional form and described the folding process as achieving a state of energy equilibrium. This highly reduced the variation among folding patterns (Go, 1983; Jerath et al., 2016). However, the number of possible protein structures resulting from the few folding patterns available for each protein still allowed an enormous number of possibilities. Computer algorithms have been developed to conduct these computations. At the same time, other efforts to reduce the possibility of tertiary and quaternary folding patterns include the consideration of the interaction between polar and nonpolar monomers or what is known as the hydrophilic or hydrophobic part of the protein. Unfortunately, computer-predicted protein folded structures are still not accurate (Díaz-Villanueva et al., 2015; Wolynes, 2005; Ben-Naim, 2012; Dill, 1995).

The assessment of the three-dimensional topology of the protein chain has also been adopted to reduce the spread of folding pattern variations. This topological method was suitably reviewed by Baker (2000). With this approach, the protein-folding is predicted using the local protein topology rather than the free-energy landscape in the inter-atomic interactions. The protein topology encompasses possible residue-residue molecular contact in a three-dimensional structure (Adhikari, 2015). Although the possible number of folding patterns can be reduced, the number of possible configurations of the residue amino acids is still high. This again resulted in a tremendous amount of computation time required to attain the right prediction of the whole protein chain to form the native structure. Notably, the use of the topology approach in solving the protein-folding mechanism has yet to yield a concessive mechanism that can be effectively used to predict protein-folding. Still, this indicates that the orientation of the residue amino acids is one of the keys to defining the protein-folding mechanism.

At this point, available evidence largely points toward that native-structure formation does not require the lowest energy state nor is random; instead, the folding is very much predefined or memorised. So, to comprehend the details of the memory, one must understand the mechanism of memorising. This brings us to the possible discovery of the true mechanism behind protein folding in the next section.

# IS THE PROTEIN STRUCTURE-FOLDING MECHANISM GOVERNED BY THE SHAPE-MEMORY EFFECT?

Many papers, when mentioning the refolding of protein chains to native structures, mostly do not report in which state the refolding is initiated. This information is very important and could have a big impact on identifying the basic concept of the protein-folding mechanism. The refolding mechanism of the tertiary and quaternary structures is far easier to understand than that of the secondary structure. This is because secondary folding involves the orientation of the residue amino acids and the folding vector of the subsequence protein molecule. Thus, most of the residual amino acids are already locked into place by the secondary structure, leaving limited possible combinations for the tertiary and quaternary (Anfinsen et al., 1961; Finkelstein & Garbuzynskiy, 2016). The successful rate of protein-chain refolding to the native structure is about 60% to 70% (Anfinsen et al., 1961). This also can be observed elsewhere (Levinthal, 1969; Levinthal, 1968; Meng & Li, 2013) using the structure-prediction algorithms.

For the formation of the secondary structure, if the Anfinsen lowest-energy hypothesis is true, then the protein-folding mechanism will become much simpler. However, the Anfinsen lowest-energy hypothesis is not consistent with the folding mechanism of protein chains in nature. From existing observations, there is no clear indication that the refolding of protein to the native primary structure occurs from the unfolded protein chain. The accuracy of this information is very crucial to establishing the concept of the protein folding mechanism. If the protein chain is capable of folding back to its native secondary structure under a suitable bio environment without the influence of RNA, this means the knowledge of folding back to the native structure exists in the protein chain itself. Portman and Takada (2001) suggest that the protein-chain folding motion is like a memory effect. This brings about the possible memory-effect phenomenon that had been observed already in many shape-memory materials. Shape-memory materials are a type of stimuli-response materials that reform to their memorised shape when the correct stimulus is applied. There is a wide range of shape-memory materials, and they respond to different stimuli such as temperature, pH, light, and magnetic and electric fields (Meng & Li, 2013). An example of memory material is the magical paperclip made from shape-memory alloys, whose shape can be altered mechanically when it is cooled; further, when it is placed in hot water, it will reform back to its original paperclip shape. Where the hot temperature is the stimulant for shape-memory alloys. This is like the protein chain that folds back to its native structure under the right bio environment, and for the protein chain, the stimulant is the right bio environment. The shape-memory mechanism at the molecular level could be different for the shape-memory alloys compared to the protein chain, but the phenomenon is the same. If this is the phenomenon that causes the protein chain to refold back to its native structure,

the protein chain would first need to undergo a memorisation process, known as shape programming. Where the native structure is programmed onto the protein chain by orientating the residue amino acids at certain vectors to direct the folding process. This process can also be assisted by the scaffolding structure where temporary folding is enacted to assist in the final folding, where the scaffolding structure will be unfolded when the native structure is achieved. This sequence of folding and unfolding creates a complex memory pathway that is referred to as the folding pathway. This had been observed by Englander and Mayne (2014; 2017), where the residue amino acid that forms the folding is termed 'fold on'. It is highly plausible for the shape-programming process to happen during the RNA binding process, which also supports the hypothesis that RNA governs the protein folding mechanism stated in the next section.

## RNA CODONS GOVERN THE PROTEIN FOLDING MECHANISM

The protein chain is proposed to be governed by a folding mechanism defined by RNA codons. Each synonymous codon for the same protein holds the protein at a specific orientation, forcing the residue amino acids to be orientated at a specific vector direction that facilitates the folding and bonding processes. For example, UUU and UUC codons both bind with the Phenylalanine (PHE) protein, but each of them will result in different folding orientations and degrees of bonding energy due to the unique bonding orientation and chemical-bonding profiles of the protein depending on the nucleotide (Frankel & Smith, 1998). The folding orientation also depends upon the next codon as different pairs of codons can result in certain folding vectors: for example, a UUC codon followed by an ACU codon and an ACU codon followed by a UUC codon will produce deferent folding orientations. Although both sequences involve combining the PHE protein and the Threonine (THR) protein, the direction of each folding orientation is different. There is a total of 64 codons made from three RNAs each, which results in a total of 8,192 ($64 \times 64 \times 2$) distinguished folding vectors. If the stop codons UAA, UAG, and UGA are not included, then the number of folding vectors is 7,442 ($61 \times 61 \times 2$). A study on *Escherichia coli* showed that the 165 known folding vectors from the database compose only 8.4% of the total possible *E. coli* folding vectors (Braselmann et al., 2013). As such, *E. coli* presents about 1,964 types of folding vectors, which is about 26.4% of the 7,442 predicted folding vectors. This result shows the folding vector of the *E. coli* is a possible sub-set of the total folding vector which is consistent with the proposed hypothesis.

Moreover, researchers have also discovered that replacing the synonymous codon in the RNA chain will lead to the onset of different protein properties, which suggests that codon alterations may result in the protein chain folding differently (Komar et al., 1999). Other investigators (Jacobson & Clark, 2016; Dykeman et al., 2014; Yu; Thommen et al., 2017; Faure et al., 2016; Sander et al., 2014; Rodnina, 2016; Carter & Wolfenden, 2015)

also observed a similar effect on the protein folding upon replacing the synonymous RNA codon in the RNA chain. In research by Zhou et al. (2015), the alteration of the synonymous codon not only changed the protein-folding structure but also affected the folding rate. This strongly supports the hypothesis that protein folding mechanisms are governed by RNA codons.

If protein-chain folding occurs according to the shape-memory effect, then there must be a process by which to imprint or program the memory onto it, which must happen during RNA binding. The chemical bonding between the nucleotide and the protein could result in a catalytic process that creates a shift in the energy landscape. Besides, depending on the physical orientation of the bonding, the catalytic binding may imprint a shifted energy landscape onto the protein molecule, causing the molecule to be programmed with said shifted energy landscape, which drives the protein to form bonds with another protein according to the specific folding vector. This catalytic process acts as the native structure-programming process for the protein chain. Possible evidence that RNA is responsible for protein folding can be found in previous research (Samanta et al., 2008), where the authors showed that the RNA-assisted folding of the protein to its native form was three to four times more successful than that of non-RNA-assisted. This suggests that the existence of RNA during protein refolding can reprogram protein chains that have lost their native shape memory and enormously increase the number of proteins able to be folded back to their native structures. This strongly supports the thought that RNA codons govern the protein-folding mechanism.

## CONCLUSION

This paper discussed the ability of proteins to fold back to their native structures without adhering to the lowest energy path and noted that random processing could be completely disproven by the shape-memory effect. Subsequently, the occurrence of a memory effect on the protein chain was described by the programming process governed by the RNA codon. Possible evidence and precursors showed strong concurrence with the protein chain shape-memory effect and supported RNA as the imprint for programming the protein chain. This escalated the hypothesis that RNA codons can govern the protein-folding mechanism. Further investigations based on all RNA codon pairs and their folding vectors are required to verify this hypothesis.

## COMPETING INTEREST

The author declares no competing interest.

# REFERENCES


Adhikari, B., Bhattacharya, D., Cao, R., & Cheng, J. (2015). CONFOLD: Residue-residue contact-guided ab initio protein folding. *Proteins: Structure, Function, and Bioinformatics*, *83* (8), 1436 – 1449. https://doi.org/10.1002/prot.24829

Alm, E. & Baker, D. (1999). Matching theory and experiment in protein folding. *Current Opinion in Structural Biology*, *9* (2), 189--196. https://doi.org/10.1016/S0959-440X(99)80027-X

Anfinsen, C. B., Haber, E., Sela, M., & White, F. H., Jr (1961). The kinetics of formation of native ribonuclease during oxidation of the reduced polypeptide chain. *Proceedings of the National Academy of Sciences of the United States of America*, *47* (9), 1309 – 1314. https://doi.org/10.1073/pnas.47.9.1309

Anfinsen, C. B. (1973). Principles that govern the folding of protein chains. *Science*, *181* (4096), 223 – 230. https://doi.org/10.1126/science.181.4096.223

Baker, D. (2000). A surprising simplicity to protein folding. *Nature*, *405*, 39 – 42. https://doi.org/10.1038/35011000

Ben-Naim, A. (2012). Levinthal's paradox revisited, and dismissed. *Open Journal of Biophysics*, *2* (2), 23 – 32. https://doi.org/10.4236/ojbiphy.2012.22004

Braselmann, E., Chaney, J. L., & Clark, P. L. (2013). Folding the proteome. *Trends in Biochemical Sciences*, *38* (7), 337 – 344. https://doi.org/10.1016/j.tibs.2013.05.001

Carter, C. W., & Wolfenden, R. (2015). tRNA acceptor stem and anticodon bases form independent codes related to protein folding. *Biophysics and Computational Biology*, *112* (24), 7489 – 7494. https://doi.org/10.1073/pnas.1507569112

Chiti, F., & Dobson, C. M. (2017). Protein misfolding, amyloid formation, and human disease: A summary of progress over the last decade. *Annual Review of Biochemistry*, *86*, 27 – 68. https://doi.org/10.1146/annurev-biochem-061516-045115

Díaz-Villanueva, J. F., Díaz-Molina, R., & García-González, V. (2015). Protein folding and mechanisms of proteostasis. *International Journal of Molecular Sciences*, *16* (8), 17193 – 17230. https://doi.org/10.3390%2Fijms160817193

Dill, K. A., Bromberg, S., Yue, K., Chan, H. S., Ftebig, K. M., Yee, David P., & Thomas, P. D. (1995). Principles of protein folding—a perspective from simple exact models. *Protein Science*, *4* (4), 561 – 602. https://doi.org/10.1002/pro.5560040401

Dobson, C. M. (2004). Principles of protein folding, misfolding and aggregation. *Seminars in Cell and Developmental Biology*, *15* (1), 3 – 16. https://doi.org/10.1016/j.semcdb.2003.12.008

Dobson, C. M. (2003). Protein folding and misfolding. *Nature*, *426*, 884 – 890. https://doi.org/10.1038/nature02261

Dykeman, E. C., Stockley, P. G., & Twarock, R. (2014). Solving a Levinthal's paradox for virus assembly identifies a unique antiviral strategy. *Biological Sciences*, *111* (14), 5361 – 5366. https://doi.org/10.1073/pnas.1319479111

Englander, S. W., & Mayne, L. (2017). The case for defined protein folding pathways. *Biophysics and Computational Biology*, *114* (31), 8253 – 8258. https://doi.org/10.1073/pnas.1706196114

Englander, S. W., & Mayne, L. (2014). The nature of protein folding pathways. *Biophysics and Computational Biology*, *111* (31), 15873 – 15880. https://doi.org/10.1073/pnas.1411798111

Faure, G., Ogurtsov, A. Y., Shabalina, S. A., & Koonin, E. V. (2016). Role of mRNA structure in the control of protein folding. *Nucleic Acids Research*, *44* (22), 10898 – 10911. https://doi.

org/10.1093/nar/gkw671
Finkelstein, A. V., Badretdin, A. J., Galzitskaya, O. V., Ivankov, D. N., Bogatyreva, N. S., Garbuzynskiy, S. O. (2017). There and back again: Two views on the protein folding puzzle. *Physics of Life Reviews*, *21*, 56 – 71. https://doi.org/10.1016/j.plrev.2017.01.025
Finkelstein, A. V., & Garbuzynskiy, S. O. (2016). Solution of Levinthal's paradox is possible at the level of the formation and assembly of protein secondary structures. *Biophysics*, *61*, 1 – 5. https://doi.org/10.1134/S0006350916010085
Frankel, A. D., & Smith, C. A. (1998). Induced folding in RNA – protein recognition: More than a simple molecular handshake. *Cell*, *92* (2), 149 – 151. https://doi.org/10.1016/S0092-8674(00)80908-3
Gō, N. (1983). Theoretical studies of protein folding. *Annual Review of Biophysics and Bioengineering*, *12*, 183 – 210. https://doi.org/10.1146/annurev.bb.12.060183.001151
Hartl, F. U., Bracher, A., & Hayer-Hartl, M. (2011). Molecular chaperones in protein folding and proteostasis. *Nature*, *475*, 324 – 332. https://doi.org/10.1038/nature10317
Honig, B. (1999). Protein folding: from the levinthal paradox to structure prediction. *Journal of Molecular Biology*, *293* (2), 283 – 293. https://doi.org/10.1006/jmbi.1999.3006
Jacobson, G. N., & Clark, P. L. (2016). Quality over quantity: Optimizing co-translational protein folding with non-'optimal'synonymous codons. *Current Opinion in Structural Biology*, *38*, 102 – 110. https://doi.org/10.1016/j.sbi.2016.06.002
Jerath, G., Hazam, P. K., Shekhar, S., & Ramakrishnan, V. (2016). Mapping the geometric evolution of protein folding motor. *PloS One*, *11* (10), e0163993. https://doi.org/10.1371/journal.pone.0163993
Karplus, M. (1997). The Levinthal paradox: Yesterday and today. *Folding and Design*, *2* (Suppl 1), S69 – S75. https://doi.org/10.1016/S1359-0278(97)00067-9
Komar, A. A., Lesnik, T., & Reiss, C. (1999). Synonymous codon substitutions affect ribosome traffic and protein folding during in vitro translation. *FEBS Letters*, *462* (3), 387 – 391. https://doi.org/10.1016/S0014-5793(99)01566-5
Kubelka, J., & Hofrichter, J., & Eaton, W. A. (2004). The protein folding 'speed limit'. Current *Opinion in Structural Biology*, *14* (1), 76 – 88. https://doi.org/10.1016/j.sbi.2004.01.013
Levinthal, C. (1968). Are there pathways for protein folding? *J. Chim. Phys.*, *65*, 44 – 45. https://doi.org/10.1051/jcp/1968650044
Levinthal, C. (1969). How to fold graciously. *Mossbaun Spectroscopy in Biological Systems Proceedings*, *67* (41), 22 – 24.
Meng, H., & Li, G. (2013). A review of stimuli-responsive shape memory polymer composites. *Polymer*, *54* (9), 2199 – 2221. https://doi.org/10.1016/j.polymer.2013.02.023
Nirenberg, M., Caskey, T., Marshall, R., Brimacombe, R., Kellogg, D., Doctor, B., Hatfield, D., Levin, J., Rottman, F., Pestka, S., Wilcox, M., & Anderson, F. (1966). The RNA code and protein synthesis. *Cold Spring Harbor Symposia on Quantitative Biology*, *31*, 11 – 24. https://doi.org/10.1101/SQB.1966.031.01.008
Onuchic, J. N., & Wolynes, P. G. (2004). Theory of protein folding. *Current Opinion in Structural Biology*, *14* (1), 70 – 75. https://doi.org/10.1016/j.sbi.2004.01.009
Portman, J. J., &Takada, S. (2001). Microscopic theory of protein folding rates. II. Local reaction coordinates and chain dynamics. *The Journal of Chemical Physics*, *114*, 5082 – 5096. https://doi.org/10.1063/1.1334663
Rodnina, M. V. (2016). The ribosome in action: Tuning of translational efficiency and protein

folding. *Protein Science*, *25* (8), 1390 – 1406. https://doi.org/10.1002/pro.2950

Samanta, D., Mukhopadhyay, D., Chowdhury, S., Ghosh, J., Pal, S., Basu, A., Bhattacharya, A., Das, A., Das, D., & DasGupta, C. (2008). Protein folding by domain V of Escherichia coli 23S rRNA: Specificity of RNA-protein interactions. *Journal of Bacteriology*, *190* (9), 3344 – 3352. https://doi.org/10.1128/JB.01800-07

Sander, I. M., Chaney, J. L., & Clark, P. L. (2014). Expanding Anfinsen's principle: Contributions of synonymous codon selection to rational protein design. *Journal of the American Chemical Society*, *136* (3), 85 – 861. https://doi.org/10.1021/ja411302m

Schuler, B., & Eaton, W. A. (2008). Protein folding studied by single-molecule FRET. *Current Opinion in Structural Biology*, *18* (1), 16 – 26. https://doi.org/10.1016/j.sbi.2007.12.003

Thirumalai, D., & Hyeon, C. (2005). RNA and protein folding: common themes and variations. *Biochemistry*, *44* (13), 4957 – 4970. https://doi.org/10.1021/bi047314+

Thommen, M., Holtkamp, W., & Rodnina, M. V. (2017). Co-translational protein folding: Progress and methods. *Current Opinion in Structural Biology*, *42*, 83 – 89. https://doi.org/10.1016/j.sbi.2016.11.020

Tompa, P., & Rose, G. D. (2011). The Levinthal paradox of the interactome. *Protein Science*, *20* (12), 2074 – 2079. https://doi.org/10.1002/pro.747

Vos, T., Allen, C., Arora, M., Barber, R. M., Bhutta, Z. A., Brown, A., Carter, A., Casey, D. C., Charlson, F. J., Chen, A. Z., Coggeshall, M., Cornaby, L., Dandona, L., Dicker, D. J., Dilegge, T., Erskine, H. E., Ferrari, A. J., Fitzmaurice, C., Flemming, T., … Tura, A. K. (2016). Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990 – 2015: A systematic analysis for the Global Burden of Disease Study 2015. *The Lancet*, *388*, 1545 – 1602. https://doi.org/10.1016/S0140-6736(16)31678-6

Waudby, C. A., Dobson, C. M., & Christodoulou, J. (2019). Nature and regulation of protein folding on the ribosome. *Trends in Biochemical Sciences*, *44* (11), 914 – 926. https://doi.org/10.1016/j.tibs.2019.06.008

Wolynes, P. G. (2006). Recent successes of the energy landscape theory of protein folding and function. *Quarterly Reviews of Biophysics*, *38* (4), 405. https://doi.org/10.1017/S0033583505004075

Young, J. C., Moarefi, I., & Hartl, F. U. (2001). Hsp90: A specialized but essential protein-folding tool. *Journal of Cell Biology*, *154* (2), 267 – 274. https://doi.org/10.1083/jcb.200104079

Yu, C. H., Dang, Y., Zhou, Z., Wu, C., Zhao, F., Sachs, M. S., & Liu, Y. (2015). Codon usage influences the local rate of translation elongation to regulate co-translational protein folding. *Molecular Cell*, 59 (5), 744 – 754. https://doi.org/10.1016/j.molcel.2015.07.018

Zhou, M., Tao, W., Fu, J., Xiao, G., & Liu, Y. (2015). Nonoptimal codon usage influences protein structure in intrinsically disordered regions. *Molecular Microbiology*, *97* (5), 974 – 987. https://doi.org/10.1111/mmi.13079

Zwanzig, R., Szabo, A., & Bagchi, B. (1992). Levinthal's paradox. *Proceedings of the National Academy of Sciences*, *89* (1), 20 – 22. https://doi.org/10.1073/pnas.89.1.20