Journal of Molecular Evolution (2020) 88:415–417
https://doi.org/10.1007/s00239-020-09948-y

EDITORIAL

Evolutionary Processes and Biophysical Mechanisms: Revisiting Why Evolved Proteins Are Marginally Stable

Amanda E. Wilson[1] · Westin M. Kosater[1] · David A. Liberles[1]

Published online: 8 May 2020

Abstract
Evolved proteins observed in natural organisms are found to be only marginally stable. Several mechanistic hypotheses have been presented to date to explain this observation. One idea that has been put forward is that active selection prevents proteins from becoming too stable to enable proper function. A second idea is that marginal stability reflects the point of mutation–selection–drift balance, where it is mutational pressure that generates marginal stability. A third idea explored in this issue of *Journal of Molecular Evolution* is that a physical limit prevents the evolution of more stable proteins rather than an evolutionary process. While the first two notions are based upon specific evolutionary processes, discussion here is aimed at reconciling evolutionary processes with the physics of protein folding, drawing upon the ideas that have been presented.

Introduction

It has been observed that proteins tend to be metastable, meaning that the free energy of folding is only slightly negative. Several explanations based upon physical and evolutionary processes have been put forward to explain this observation. In this issue of *Journal of Molecular Evolution* (Martin and Vila 2020), it is hypothesized that the upper bound limit for marginal stability in proteins is a universal phenomenon derived from the physical nature of folding specificity, and is not affected by evolutionary processes or the nature of chemical posttranslational modifications (Vila 2019; Martin and Vila 2020). In contrast, two hypotheses have been previously presented based upon distinct evolutionary processes. One evolutionary argument is that metastability is an equilibrium point in mutation–selection–drift balance (Taverna and Goldstein 2002; Goldstein 2013). A very different evolutionary argument is that more stable proteins would not be particularly rare in sequence space and that strong selection for proper function generates metastability (DePristo et al. 2005). These cases and potential reconciliations between them are presented here.

Observations of the Physics of Natural Proteins

A recent paper in *Journal of Molecular Evolution* (Martin and Vila 2020) builds upon a prior publication (Vila 2019) and makes several observations about the nature of protein stability. The first observation is that large multi-molecular complexes, like the ribosome, and smaller enzyme structures have similar stabilities, which rests upon characterizations of proteomes as being broadly marginally stable (Ghosh and Dill 2010). Based upon a conceptual model of protein folding that contrasts the sets of inter-atomic contacts and their fluctuations that would be possible in native and non-native structures, a heuristic weak relationship between molecular weight and folding free energy is proposed. Together with the small observed effects of point mutations on protein stability that is part of this observation, an upper bound to the stability of proteins is proposed. The distribution of effects on protein stability of observed point mutations is biased toward those of small effect that maintain protein marginal stability and would likely be evolutionarily neutral. While the case is presented as a physical limit to protein stability, evolutionary processes that may contribute to this observed limit and those that result from it are described (Martin and Vila 2020) and will be explored further here.

✉ David A. Liberles
daliberles@temple.edu

[1] Department of Biology and Center for Computational Genetics and Genomics, Temple University, Philadelphia, PA 19122, USA

The Physical Limits of Folding Stability

While there appears to be a physical limit to the stability of naturally observed folded proteins, it is known from protein design and computational analysis that extremely stable protein structures can be generated (discussed in Goldstein 2011, 2013). Further, more stable proteins can evolve in thermophilic organisms that live at high temperatures, using an increased hydrophobic content in proteins that is mediated through the genetic code by amino acid mutations from altered nucleotide usage patterns (Goldstein 2007; Zeldovich et al. 2007). Without the changed nucleotide usage pattern, mutations giving rise to more stable proteins may be rare because combinations of sequences with mesophilic amino acid usage that are extremely stable would be rare and improbable to sample (Zeldovich et al. 2007, see Lau and Dill 1989 for a lattice model case). Further, it should be noted that while kinetic control of protein folding is possible, lattice models of proteins that initially fold with kinetic control were observed to evolve native states that frequently were the thermodynamic minimum in the absence of a selective pressure to retain kinetic control (Govindarajan and Goldstein 1998).

Does Selection Give Rise to Marginal Stability?

Against this backdrop, it should be noted that proteins do not just fold, they also need to function, not aggregate, and behave under a number of specific constraints (DePristo et al. 2005; Chi and Liberles 2016). While these additional constraints do act and will restrict sequence space (Liberles et al. 2011), they probably do not require substantial stability beyond metastable observations in most circumstances. Prevention of self-aggregation for example might lead to a reduced selective pressure for hydrophobic residues that could provide additional stability (DePristo et al. 2005). Further, it has been argued that proper function and eventual protein degradation by proteases require that protein stabilities have a selective upper bound (DePristo et al. 2005). While it is clear that global folding stability is an over-simplification of the constraints on a folded protein, one can ask the question if such selection is necessary, given the improbability that sets of mutations would give rise to very stable proteins.

A Case for Mutation–Selection–Drift Balance

The efficacy of selection to reduce deleterious mutations and bring beneficial mutations to fixation is predicted to be affected by the mutation rate and effect distribution, the strength of selection, and the amount of genetic drift modulated by effective population size (N_e). The equilibrium between these forces is referred to as mutation–selection–drift balance. In the context of protein stability, the equilibrium is the point where the effects of destabilizing mutations on a protein are compensated by the effect of selection favoring stabilizing mutations (Goldstein 2011). Simulations have shown populations with an N_e of 10^4, equilibrium resulted in proteins with ΔG of approximately -7 kcal mol^{-1}, while populations with an N_e of 10^8 equilibrated at ΔG of -12 kcal mol^{-1} (Goldstein 2013). It was noted that these simulated results have a similar ΔG to proteins produced in nature, but hill-climbing algorithms find sequences an order of magnitude more stable that maintain functional activity, also consistent with experimental observations (Goldstein 2011). Higher effective population sizes lead to more stable proteins because selection is more efficient, but is not orders of magnitude more stable because the mutations on more stable proteins have a smaller effect. Therefore, selection cannot counteract the distribution of mutational effects on stable proteins in large effective populations. These results support the notion that metastability of proteins is the equilibrium point of mutation, selection and genetic drift under constant effective population size. The role of effective population in the dynamics of protein stability evolution is complicated and our current understanding of this depends upon modeling assumptions that have been made both about how proteins evolve and about the evolutionary history of effective population sizes on inter-specific timescales. However, it is apparent that mutational pressure from the distribution of protein stabilities as densities in sequence space is sufficient to give rise to marginal stability.

Concluding Thoughts

It has become established that naturally observed proteins are marginally stable. The root causes of this rest in the link between the physical nature of protein structure and stability and the population genetic and evolutionary processes that play out on top of this, resulting in both mutational pressure and selective pressures. In this issue of *Journal of Molecular Evolution*, Martin and Vila (2020) have added one more piece to an ongoing discussion about these processes and mechanisms.

Acknowledgements We thank Darius Balciunas, Sarah Emel, Richard Goldstein, Dan Weinreich, and Jorge Vila for reading an earlier draft of this piece and providing comments.

References

Chi PB, Liberles DA (2016) Selection on protein structure, interaction, and sequence. Protein Sci 25(7):1168–1178. https://doi.org/10.1002/pro.2886

DePristo MA, Weinreich DM, Hartl DL (2005) Missense meanderings in sequence space: a biophysical view of protein evolution. Nat Rev Genet 6(9):678–687

Ghosh K, Dill K (2010) Cellular proteomes have broad distributions of protein stability. Biophys J 99(12):3996–4002. https://doi.org/10.1016/j.bpj.2010.10.036

Goldstein RA (2007) Amino-acid interactions in psychrophiles, mesophiles, thermophiles, and hyperthermophiles: insights from the quasi-chemical approximation. Protein Sci 16(9):1887–1895

Goldstein RA (2011) The evolution and evolutionary consequences of marginal thermostability in proteins. Proteins 79(5):1396–1407. https://doi.org/10.1002/prot.22964

Goldstein RA (2013) Population size dependence of fitness effect distribution and substitution rate probed by biophysical model of protein thermostability. Genome Biol Evol 5(9):1584–1593. https://doi.org/10.1093/gbe/evt110

Govindarajan S, Goldstein RA (1998) On the thermodynamic hypothesis of protein folding. Proc Natl Acad Sci USA 95(10):5545–5549

Lau KF, Dill KA (1989) A lattice statistical mechanics model of the conformational and sequence spaces of proteins. Macromolecules 22(10):3986–3997

Liberles DA, Tisdell MD, Grahnen JA (2011) Binding constraints on the evolution of enzymes and signalling proteins: the important role of negative pleiotropy. Proc Biol Sci 278(1714):1930–1935. https://doi.org/10.1098/rspb.2010.2637

Martin OA, Vila JA (2020) The marginal stability of proteins: how the jiggling and wiggling of atoms is connected to neutral evolution. J Mol Evol. https://doi.org/10.1007/s00239-020-09940-6

Taverna DM, Goldstein RA (2002) Why are proteins marginally stable? Proteins 46(1):105–109

Vila JA (2019) Forecasting the upper bound free energy difference between protein native-like structures. Physics A 533:122053

Zeldovich KB, Berezovsky IN, Shakhnovich EI (2007) Protein and DNA sequence determinants of thermophilic adaptation. PLoS Comput Biol 3(1):e5