Proteins' Evolution upon Point Mutations

Jorge A. Vila*

Read Online

ACCESS | Metrics & More | Article Recommendations

ABSTRACT: As the reader must be already aware, state-of-the-art protein folding prediction methods have reached a smashing success in their goal of accurately determining the three-dimensional structures of proteins. Yet, a solution to simple problems such as the effects of protein point mutations on their (i) native conformation; (ii) marginal stability; (iii) ensemble of high-energy nativelike conformations; and (iv) metamorphism propensity and, hence, their evolvability, remains as an unsolved problem. As a plausible solution to the latter, some properties of the amide hydrogen-deuterium exchange, a highly sensitive probe of the structure, stability, and folding of proteins, are assessed from a new perspective. The preliminary results indicate that the protein marginal stability change upon point mutations provides the

necessary and sufficient information to estimate, through a Boltzmann factor, the evolution of the amide hydrogen exchange protection factors and, consequently, that of the ensemble of folded conformations coexisting with the native state. This work contributes to our general understanding of the effects of point mutations on proteins and may spur significant progress in our efforts to develop methods to determine the appearance of new folds and functions accurately.
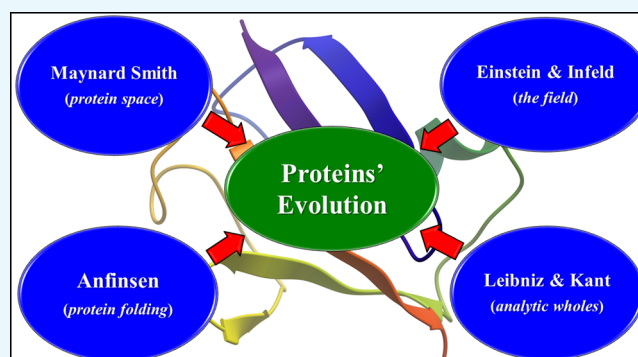
INTRODUCTION

How proteins' point mutations impact their evolvability is of paramount importance in biology, from molecular evolution to structural biology; for example, nuclear magnetic resonance results reveal that "proteins adopt unstable, high-energy states that exist for fractions of a second but can have key biological roles".[1] The long-standing question is how to resolve this problem without considering epistasis effects explicitly. Before we embark on unraveling a possible solution to this enigma, it is worth noting that the term epistasis has been used with various analogous meanings, although it is commonly defined as a phenomenon that "... occurs when the combined effect of two or more mutations differs from the sum of their individual effects...".[2] We should point out here a remarkable equivalence between this definition of epistasis with Leibniz and Kant's notion of space (and time) devised as "analytic wholes", i.e., the one where "...its priority makes it impossible to obtain it by the additive synthesis of previously existing entities...".[3] Beyond this philosophical thought, the following question arises, why should we be interested in epistasis? The main reason is that epistasis could have a remarkable impact on the evolution of proteins by either restricting their trajectories or opening new paths to new sequences that would otherwise have been inaccessible.[4−6] However, despite the simplicity of the definition of epistasis and the colossal progress in the prediction of protein structures, the answers to simple questions remain elusive because "...there is currently no means to predict specific epistasis from a protein

sequence or structure...".[2] What if we analyze a far simpler problem? For example, can point mutation effects be forecast accurately? Unfortunately, the answer still is no. To determine the nature of this problem, which includes that of epistasis, we should start by identifying the molecular origin and the main factors affecting point mutations. In this regard, it is not enough to consider the protein sequence nor the mutation-types—it could also be a post-translational modification[7]—but a precise determination of the "field" between and around amino acids. The relevance of the "field" for an accurate description of any physical problem was highlighted by Einstein and Infeld[8] in the following terms: "A new concept appears in physics, the most important invention since Newton's time: the field. It needed great scientist imagination to realize that it is not the charges nor the particles, but the field in the space between the charges and the particles which is essential for the description of physical phenomena." Application of this concept on structural biology started with the pioneer development of all-atom "force-field"[9−13] aimed to predict the three-dimensional structure of proteins with the only knowledge of the amino-acid sequence—

the protein folding problem.[14] However, a definitive solution to this problem has been elusive since then, e.g., how a sequence encodes the protein folding remains unknown,[15] even though the protein's three-dimensional structure can be accurately determined.[16] Consequently, and beyond any doubt, an accurate determination of point mutation effects is an unsolved problem still.[17,18] Indeed, a large number of methods and approaches used to predict protein stability upon point mutation, e.g., by using physical, statistical, or empirical "force-field", respectively, or machine learning methods,[19,20] show limited performance and suffer from caveats.[19,21−25] Therefore, we should focus on the global rather than on the specific mutation effects.

In contrast to the mutation-specific effects, the global distribution of the proteins' stability upon mutations[26,27] can be forecast with acceptable accuracy by a bi-Gaussian function.[21] Such changes in the proteins' stability due to point mutations can be determined experimentally from the unfolding Gibbs free energy (ΔG_U) between the wild-type (wt) and the mutant (m) protein, viz., as $\Delta\Delta G_U = (\Delta G_U^m - \Delta G_U^{wt})$.[28] Consequently, considering that point mutations mainly affect the native-state stability,[29] the observed change on the unfolding Gibbs free energy ($\Delta\Delta G_U$) should represent, fundamentally, the change ($\Delta\Delta G$) in the protein marginal stability (ΔG), which refers to the Gibbs free-energy gap between the native state and the first unfolded state.[7,30,31] Let us provide some pieces of evidence that support this important conjecture. The proteins' free energy of unfolding (ΔG_U) spans a wide range of variations, viz., between 5 and 25 kcal/mol,[32] revealing the complexity of the "protein folding problem".[14,33−35] However, its range of variation upon point mutations ($\Delta\Delta G_U$) is small and well-defined, revealing the validity of the thermodynamic hypothesis or Anfinsen dogma,[14] as explained next. The absolute values of the unfolding Gibbs free energy changes ($|\Delta\Delta G_U|$)—from the histogram of more than 5200 point mutation data obtained by using urea and thermal unfolding experiments[29]—are within the following narrow range of variation: $|\Delta\Delta G_U| \leq \sim7.5$ kcal/mol. Notably, this boundary value for $|\Delta\Delta G_U|$ conforms with the proteins' marginal stability upper bound limit, namely, ~7.4 kcal/mol,[31] which (i) is a universal feature of proteins, i.e., obtained regardless of the fold-class or its amino-acid sequence;[31] (ii) is a consequence of Anfinsen's dogma validity;[31,36] and (iii) represents a threshold beyond which a conformation will unfold and become nonfunctional.[7,36,37] The latter means that changes in the Gibbs free-energy gap size ($\Delta\Delta G$) between the native state and the first unfolded state cannot be larger than ~7.4 kcal/mol, e.g., as it occurs for the single mutants of the green fluorescent protein from *Aequorea victoria* that loses $\sim100\%$ of the log fluorescence (native function) if $\Delta\Delta G_U > \sim7.5$ kcal/mol.[38] Consequently, assuming $\Delta\Delta G \sim \Delta\Delta G_U$—with the latter being experimentally determined—is a reasonable strategy to obtain a reliable assessment of the change on the protein marginal stability upon point mutations and, from here, their effects on the ensemble of folded conformations coexisting with the native state,[36] as shown later.

The gained knowledge on protein (i) stability;[21,29,31,37,39−41] (ii) metamorphism, characterized by the existence of two or more folds with a significant structural difference between them;[37,42−51] and (iii) evolvability, the ability of a biological system to provide, by mutation and selection, phenotypic variation, has been enormous.[50,52−58] This will enable us to examine below, in light of evolution, how point mutations could impact each of those issues.

■ RESULTS AND DISCUSSION

Single-Point Mutation Effects. The unfolding Gibbs free energy changes upon mutation ($\Delta\Delta G_U$), in kcal/mol, instantly enable us to determine if they are positive (stabilizing) or negative (destabilizing) contributions to the protein's marginal stability. However, their impact on both the ensemble of folded conformations in equilibrium with the native state or the metamorphism propensity cannot be straightforwardly inferred. To solve this issue, the amide hydrogen exchange (HX) may be used because it is a sensitive probe to assess changes in the protein native-state structure.[59−66] Indeed, their use could bring precise information on the structural changes that could occur upon mutations and, consequently, on their impact on the protein marginal stability.[36] This is possible because the intra- and intermolecular hydrogen bonds are dependent on the protein native-state structure and the milieu. A simple example will be enough to illustrate this methodology. Shirley et al.[67] accurately determined the urea and thermal unfolding average free-energy change ($\Delta\Delta G_U$) on ribonuclease T1 (Rnase T1) for 12-point mutations involving Tyr → Phe, Ser → Ala, and Asn → Ala amino acids, respectively. As a result, the observed destabilizing average $\Delta\Delta G_U$ values were within the following range of variation: ≈-0.5 kcal/mol (for Tyr57 → Phe) to ~-2.9 kcal/mol (for Asn81 → Ala). Before we proceed, let us remember the following: first, $\Delta\Delta G \sim \Delta\Delta G_U$ will provide us with the Gibbs free energy change in the protein's marginal stability upon point mutation; second, the amide HX protection factor (P_f) for a protein in their native state, i.e., in the EX2 limit,[68] is given by the following equation $\Delta G_{HX} = RT \ln P_f$,[68] where ΔG_{HX} represents the Gibbs free-energy change for the opening/closing equilibrium,[64,68] R is the gas constant, and T is the absolute temperature. Because our interest focuses on a particular region of the conformational space, namely, in the Boltzmann ensemble of folded states in equilibrium with the native state, the following relation $\Delta G_{HX} \sim \Delta G$ should hold.[36] Consequently, upon a point mutation, the following relations $\Delta\Delta G_{HX m,wt} \sim \Delta\Delta G = (\Delta G_m - \Delta G_{wt}) = RT \ln (P_{f,m}/P_{f,wt})$, where $P_{f,m}$ and $P_{f,wt}$ are the corresponding protection factors for the mutant (m) and the wild-type (wt) protein, respectively, should also hold. Therefore, if $\Delta\Delta G \sim -2.9$ kcal/mol, then $P_{f,m} \sim P_{f,wt} \times 10^{-2}$, where $P_{f,wt}$ represents the resistance of the amide HX in the wild-type native state relative to that of the highest free-energy conformation in the ensemble of folded states.[36] In other words, because Asn81 → Ala is a destabilizing mutation, it will leave a native state for the mutant (m) that is ~100 times less resistant to the amide HX than that of the wild-type protein (wt). Therefore, the point mutations change not only the stability of the native state[29] but also the structural dispersion in the ensemble of folded conformations coexisting with it. This conjecture is in line with convincing theoretical simulations of the HX mechanism on proteins.[69,70] Indeed, such simulations show that sizeable structural differences—in the ensemble of folded conformations relative to the native state—are not only likely but necessary for accurately analyzing the observed amide HX. From this point of view, it is reasonable to assume that a point mutation will introduce structural/energetic fluctuations in the ensemble of native folds in equilibrium with the native state (see Figure 1). The latter would be of great impact on protein evolution since the existing piece of evidence shows that small changes in protein structure are essential to their function.[1] If the protein were a metamorphic one—an attribute encoded in its amino-acid sequence—a point mutation could
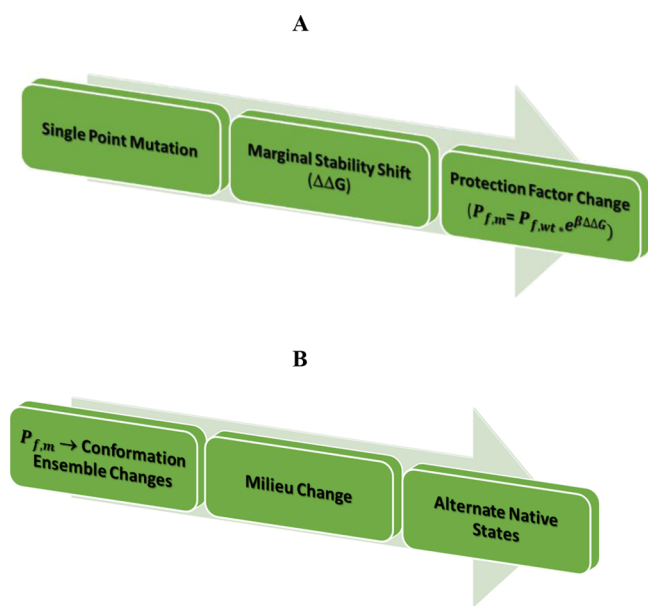
A

B

Figure 1. Easy charts of the single-point mutation effects in terms of the protection factor (P_f), representing the resistance of the amide HX in the native state relative to that of the highest free-energy conformation in the ensemble of folded states, and the protein marginal-stability change ($\Delta\Delta G$). (A) Protection factor change for the mutant ($P_{f,m}$) differs, from one of the wild-types, by a Boltzmann factor that is a function of $\Delta\Delta G$. (B) Changes in the ensemble of conformations coexisting with the native state enable, e.g., upon a slight change in the milieu, the appearance of alternate native states, as for the metamorphic proteins and, hence, the occurrence of new functions.

modify, a priori, its metamorphism propensity. Regardless of this, the appearance of fluctuations in the ensemble of native folds, a change in the milieu or both, could allow redistribution of their folded state ratio[36,37]—determined by its Boltzmann factors—and, hence, could benefit/impair the thermodynamic equilibrium between highly dissimilar (metamorphic) folded states. This could be of paramount importance to identify critical amino acids for the arise—or disappearance—of metamorphism in proteins such as in the study of the appearance of new folds and functions upon a mutation.[49,50] In addition to all the above, the protein evolvability should also be affected by point mutation because it is well known that stability promotes it.[53,54,71,72]

Mutation Effects in Light of Evolution. We are now in good condition to determine how a series of point mutations will affect the protein's marginal-stability change. To solve this, let us frame the problem within the protein space model where "...if evolution by natural selection is to occur, functional proteins must form a continuous network, which can be traversed by unit mutational steps without passing through nonfunctional intermediates...".[73] Implicit in this modeling is that any functional protein that pertains to the protein space obeys Anfinsen's dogma.[36,37] Then, a walk in that protein space enables us to determine, after j consecutive point mutation steps, the following relations of interest:

$$(\Delta G_U^1 - \Delta G_U^{wt}) + \sum_{s=2}^{j} \Delta\Delta G_U^s$$
$$= (\Delta G_U^j - \Delta G_U^{wt})$$
$$= \Delta\Delta G_U^j \qquad (1)$$

where $\Delta\Delta G_U^s = (\Delta G_U^s - \Delta G_U^{s-1})$ for $s > 1$, and wt represents the wild type. Then, considering

$$(\Delta G^1 - \Delta G^{wt}) + \sum_{s=2}^{j} \Delta\Delta G^s = (\Delta G^j - \Delta G^{wt}) = \Delta\Delta G^j$$

$$= RT\ln\left(\frac{P_{f,j}}{P_{f,wt}}\right), \quad \text{and} \quad \Delta\Delta G^j \sim \Delta\Delta G_U^j. \qquad (2)$$

the following relationship, in terms of the observable $\Delta\Delta G_U^j$, can be obtained:

$$P_{f,j} \sim P_{f,wt} \, e^{\beta\Delta\Delta G_U^j} \qquad (3)$$

with $\beta = 1/RT$, and $P_{f,j}$ and $P_{f,wt}$ are the corresponding factors for the protein after j point mutation steps and the wild-type (wt) native state, respectively (as shown in Figure 1 for $j = 1$). In other words, eq 3 represents, after navigating the protein space as an abstract model of evolution, the changes on both the wild-type protein native-state stability ($\Delta\Delta G_U^j$) and the ensemble of folded conformations coexisting with it ($P_{f,j}$). It is worth noting that the results in eq 1−3 are valid even if there are k out of j mutations (with $k < j - 1$) leading to nonfunctional proteins, e.g., when a mutation leads to a free-energy change ($\Delta\Delta G$) larger than a marginal stability upper bound threshold of ∼7.4 kcal/mol.[31,36] Consideration of this problem is relevant for two reasons: first, because most of the point mutations are destabilizing[29,55,72,74] and second because the evolutionary trajectories in the protein sequence space are assumed to be reversible although the genotypic irreversibly should not be dismissed.[75]

Let us assume that the ancestor and target protein sequences, respectively, are kept fixed during the evolutive process, as in the word game.[73] In this game, one word changes into another by replacing one letter at a time, e.g., transforming the word "NCSI" to "IRNL",[76] where each letter represents an amino acid in the single-letter code. Then, according to eq 3, turning "NCSI" into "IRNL" will lead to the same result, in terms of $P_{f,j}$, whatever the mutational trajectories linking these two words are. That is feasible because $\Delta\Delta G^j$ as $\Delta\Delta G_U^j$ is a state function. Thus, nature follows any evolutive path if there is no penalty for doing so. In actual applications, the latter means that some trajectories could have a small (or null) chance of realization.[77] The word game also enables us to rationalize the complexity of the protein evolution analysis, e.g., in terms of either evolutionary trajectories or marginal stability changes. Let us briefly discuss the pros and cons of each of these approaches.

On the one hand, an analysis of the protein evolution in terms of the evolutionary trajectories implies knowing with precision, in each step, the letter to be mutated (amino-acid identity), the background where the mutation occurs (sequence and milieu), and the epistasis effects that could take place. An accurate solution to this problem is a daunting task,[78] although of great practical relevance, e.g., how to turn a protein to exhibit the desired function—as it happens on directed evolution applications.[54,72] A solution to this problem is exacerbated by the fact that neutral mutations, aside from epistasis effects, also need to be considered because they may play a critical role in the transition from one amino acid to another.[53,71,79−81] In other words, neutral mutations (which are invisible to natural selection) may compensate for the effects of destabilizing mutations though beneficial from the functional point of view.[72,80,81]

C

On the other hand, if protein stability is an essential factor for evolution,[74,82] then it is reasonable to think about the whole (marginal-stability evolution) rather than the parts (evolutionary-trajectories) because the former involves the latter and, consequently, all factors affecting protein evolution—including specific, nonspecific, or "high-order" epistasis effects.[4,5,78,83] Yet, we should keep in mind that Leibniz and Kant's notion of space (and time),[3] devised as "analytic wholes", highlights that the whole is more than the sum of the parts, although this does not imply the irrelevance of the latter. Indeed, as noted above for the directed evolution applications, if our interest focuses on understanding why nature adopts one among all possible evolutionary trajectories,[78] detailed consideration of the epistasis effects would be unavoidable.

CONCLUSIONS

The above analysis stresses the importance of the protein marginal-stability change analysis upon point mutations not solely because it appears as one of the main factors governing protein evolution but also because it provides a straightforward path to estimate—without considering epistasis effects explicitly—the fluctuations that may occur in the Boltzmann ensemble of folded conformations in equilibrium with its native state and, hence, their possible impact on the progress toward new architectures and functions.

AUTHOR INFORMATION

Corresponding Author

Jorge A. Vila − *IMASL-CONICET, Universidad Nacional de San Luis, 5700 San Luis, Argentina;* orcid.org/0000-0001-7557-9350; Email: jorgevila84@gmail.com

Complete contact information is available at:
https://pubs.acs.org/10.1021/acsomega.2c01407

Notes
The author declares no competing financial interest.

ACKNOWLEDGMENTS

DEDICATION

Dedicated in memoriam to Harold A. Scheraga, a highly respected scientist in the protein science field.

REFERENCES

(1) Stiller, J. B.; et al. Structure determination of high-energy states in a dynamic protein ensemble. *Nature* **2022**, *603*, 528−535.

(2) Miton, C. M.; Chen, J. Z.; Ost, K.; Anderson, D. W.; Tokuriki, N. Statistical analysis of mutational epistasis to reveal intramolecular interaction networks in proteins. *Methods Enzymol.* **2020**, *643*, 243−280.

(3) Gómez, R. J. *Leibniz's Spark of Kant's Great Light, chapter 14 in 'Thought, Language, and Ontology'*; Orilia, F.; Rapaport, W., Eds.; Kluwer Academic Publishers: Netherlands, 1998; pp 313−329.

(4) Miton, C. M.; Tokuriki, N. How mutational epistasis impairs predictability in protein evolution and design. *Protein Sci.* **2016**, *25*, 1260−1272.

(5) Domingo, J.; Baeza-Centurion, P.; Lehner, B. The Causes and Consequences of Genetic Interactions (Epistasis). *Annu. Rev. Genomics Hum. Genet.* **2019**, *20*, 433−460.

(6) Miton, C. M.; Buda, K.; Tokuriki, N. Epistasis and intramolecular networks in protein evolution. *Curr. Opin. Struct. Biol.* **2021**, *69*, 160−168.

(7) Martin, A. O.; Vila, J. A. The Marginal Stability of Proteins: How the Jiggling and Wiggling of Atoms is Connected to Neutral Evolution. *J. Mol. Evol.* **2020**, *88*, 424−426.

(8) Einstein, A.; Infeld, L. *The Evolution of Physics*; Simon and Schuster, Inc.: New York, 1961, pp 244.

(9) Némethy, G.; Scheraga, H. A. Theoretical determination of sterically allowed conformations of a polypeptide chain by a computer method. *Biopolymers* **1965**, *3*, 155−184.

(10) Gibson, K. D.; Scheraga, H. A. Minimization of polypeptide energy. I. Preliminary structures of bovine pancreatic ribonuclease S-peptide. *Proc. Natl. Acad. Sci. U. S. A.* **1967**, *58*, 420−427.

(11) Scheraga, H. A. Calculations of conformations of polypeptides. *Adv. Phys. Org. Chem.* **1968**, *6*, 103−184.

(12) Lifson, S.; Warshel, A. Consistent Force Field for Calculations of Conformations, Vibrational Spectra, and Enthalpies of Cycloalkane and n-Alkane Molecules. *J. Chem. Phys.* **1968**, *49*, 5116−5129.

(13) Momany, F. A.; McGuire, R. F.; Burgess, A. W.; Scheraga, H. A. Energy parameters in polypeptides. VII. Geometric parameters, partial atomic charges, nonbonded interactions, hydrogen bond interactions, and intrinsic torsional potentials for the naturally occurring amino acids. *J. Phys. Chem.* **1975**, *79*, 2361−2381.

(14) Anfinsen, C. B. Principles that govern the folding of protein chains. *Science* **1973**, *181*, 223−230.

(15) Cramer, P. AlphaFold2 and the future of structural biology. *Nat. Struct. Mol. Biol.* **2021**, *28*, 704−705.

(16) John, J.; et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **2021**, *596*, 583−589.

(17) Serpell, L. C.; Radford, S. E.; Otzen, D. E. AlphaFold: A Special Issue and A Special Time for Protein Science. *J. Mol. Biol.* **2021**, *433*, No. 167231.

(18) Buel, G. R.; Walters, K. J. Can AlphaFold2 predict the impact of missense mutations on structure? *Nat. Struct. Mol. Biol.* **2022**, *29*, 1−2.

(19) Khan, S.; Vihinen, M. Performance of protein stability predictors. *Hum. Mutat.* **2010**, *31*, 675−684.

(20) Sanavia, T.; Birolo, G.; Montanucci, L.; Turina, P.; Capriotti, E.; Fariselli, P. Limitations and challenges in protein stability prediction upon genome variations: towards future applications in precision medicine. *Comput. Struct. Biotechnol. J.* **2020**, *18*, 1968−1979.

(21) Tokuriki, N.; Stricher, F.; Schymkowitz, J.; Serrano, L.; Tawfik, D. S. The Stability Effects of Protein Mutations Appear to be Universally Distributed. *J. Mol. Biol.* **2007**, *369*, 1318−1332.

(22) Potapov, V.; Cohen, M.; Schreiber, G. Assessing computational methods for predicting protein stability upon mutation: good on average but not in the details. *Protein Eng., Des. Sel.* **2009**, *22*, 553−560.

(23) Kulshreshtha, S.; Chaudhary, V.; Goswami, G. K.; Mathur, N. Computational approaches for predicting mutant protein stability. *J. Comput.-Aided Mol. Des.* **2016**, *30*, 401−412.

(24) Caldararu, O.; Blundell, T. L.; Kepp, K. P. A base measure of precision for protein stability predictors: structural sensitivity. *BMC Bioinf.* **2021**, *22*, 2−14.

(25) Pancotti, C.; Benevenuta, S.; Birolo, G.; Alberini, V.; Repetto, V.; Sanavia, T.; Capriotti, E.; Fariselli, P. Predicting protein stability changes upon single-point mutation: a thorough comparison of the available tools on a new dataset. *Briefings Bioinf.* **2022**, *23*, No. bbab555.

(26) Guerois, R.; Nielsen, J. E.; Serrano, L. Predicting Changes in the Stability of Proteins and Protein Complexes: A Study of More Than 1000 Mutations. *J. Mol. Biol.* **2022**, *320*, 369−387.

(27) Schymkowitz, J. W.; Rousseau, F.; Martins, I. C.; Ferkinghoff-Borg, J.; Stricher, F.; Serrano, L. Prediction of water and metal binding sites and their affinities by using the Fold-X force field. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *2005*, 10147−10152.

(28) Bigman, L. S.; Levy, Y. Stability Effects of Protein Mutations: The Role of Long-Range Contacts. *J. Phys. Chem. B* **2018**, *122*, 11450−11459.

(29) Zeldovich, K. B.; Chen, P.; Shakhnovich, E. I. Protein stability imposes limits on organism complexity and speed of molecular evolution. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 16152−16157.

(30) Hormoz, S. Amino acid composition of proteins reduces deleterious impact of mutations. *Sci. Rep.* **2013**, *3*, 1−10.

(31) Vila, J. A. Forecasting the upper bound free energy difference between protein native-like structures. *Phys. A* **2019**, *533*, No. 122053.

(32) Gromiha, M. M.; Anoosha, P.; Huang, L. T. Applications of Protein Thermodynamic Database for Understanding Protein Mutant Stability and Designing Stable Mutants. *Methods Mol. Biol.* **2016**, *1415*, 71−89.

(33) Levinthal, C. Are There Pathways for Protein Folding? *J. Chim. Phys.* **1968**, *65*, 44−45.

(34) Anfinsen, C. B.; Scheraga, H. A. Experimental and Theoretical Aspects of Protein Folding. *Adv. Protein Chem.* **1975**, *29*, 205−300.

(35) Dill, K. A.; Ozkan, B. S.; Shell, M. S.; Weikl, T. R. The protein folding problem. *Annu. Rev. Biophys.* **2008**, *37*, 289−316.

(36) Vila, J. A. Thoughts on the protein's native state. *J. Phys. Chem. Lett.* **2021**, *12*, 5963−5966.

(37) Vila, J. A. Metamorphic proteins in light of Anfinsen's dogma. *J. Phys. Chem. Lett.* **2020**, *11*, 4998−4999.

(38) Sarkisyan, K. S.; Bolotin, D. A.; Meer, M. V.; et al. Local fitness landscape of the green fluorescent protein. *Nature* **2016**, *533*, 397−401.

(39) Albert, T. Mutational effects on protein stability. *Annu. Rev. Biochem.* **1989**, *58*, 765−792.

(40) Taverna, D. M.; Goldstein, R. A. Why Are Proteins Marginally Stable? *Proteins* **2002**, *46*, 105−109.

(41) Williams, P. D.; Pollock, D. D.; Goldstein, R. A. Functionality and the evolution of marginal stability in proteins: Inferences from lattice simulations. *Evol. Bioinf. Online* **2006**, *2*, 91−101.

(42) Luo, X.; Yu, H. Protein Metamorphosis: The two-state behavior of Mad2. *Structure* **2008**, *16*, 1616−1625.

(43) Murzin, A. G. Biochemistry. Metamorphic proteins. *Science* **2008**, *320*, 1725−1726.

(44) Tuinstra, R. L.; Peterson, F. C.; Kutlesa, S.; Elgin, E. S.; Kron, M. A.; Volkman, B. F. Interconversion between two unrelated protein folds in the lymphotactin native state. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 5057−5062.

(45) Alexander, P. A.; He, Y.; Chen, Y.; Orban, J.; Bryan, P. N. A minimal sequence code for switching protein structure and function. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *2009*, 21149−21154.

(46) Yadid, I.; Kirshenbaum, N.; Sharon, M.; Dym, O.; Tawfik, D. S. Metamorphic proteins mediate evolutionary transitions of structure. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 7287−7292.

(47) Lella, M.; Mahalakshmi, R. Metamorphic Proteins: Emergence of Dual Protein Folds from One Primary Sequence. *Biochemistry* **2017**, *56*, 2971−2984.

(48) Dishman, A. F.; Volkman, B. F. Unfolding the Mysteries of Protein Metamorphosis. *ACS Chem. Biol.* **2018**, *13*, 1438−1446.

(49) Shortle, D. One sequence plus one mutation equals two folds. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 21011−21012.

(50) Dishman, A. F.; Tyler, R. C.; Fox, J.; Kleist, A. B.; Prehoda, K. E.; Babu, M. M.; Peterson, F. C.; Volkman, B. F. Evolution of fold switching in a metamorphic protein. *Science* **2021**, *371*, 86−90.

(51) Avadhani, V. S.; Mondal, S.; Banerjee, S. Mapping Protein Structural Evolution upon Unfolding. *Biochemistry* **2022**, *61*, 303−309.

(52) James, L. C.; Tawfik, D. S. Conformational diversity and protein evolution -a 60-year-old hypothesis revisited. *Trends Biochem. Sci.* **2003**, *28*, 361−368.

(53) Bloom, J. D.; Labthavikul, S. T.; Otey, C. R.; Arnold, F. H. Protein stability promotes evolvability. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 5869−5874.

(54) Romero, P. A.; Arnold, F. H. Exploring protein fitness landscapes by directed evolution. *Nat. Rev. Mol. Cell. Biol.* **2009**, *10*, 866−876.

(55) Tokuriki, N.; Tawfik, D. S. Protein dynamics and evolvability. *Science* **2009**, *324*, 203−207.

(56) Tokuriki, N.; Tawfik, D. S. Stability effects of mutations and protein evolvability. *Curr. Opin. Struct. Biol.* **2009**, *19*, 596−604.

(57) Breen, M.; Kemena, C.; Vlasov, P.; et al. Epistasis as the primary factor in molecular evolution. *Nature* **2012**, *490*, 535−538.

(58) Goldstein, R. A. The evolution and evolutionary consequences of marginal thermostability in proteins. *Proteins* **2018**, *79*, 1396−1407.

(59) Hvidt, A.; Linderstrøm-Lang, K. Exchange of hydrogen atoms in insulin with deuterium atoms in aqueous solutions. *Biochim. Biophys. Acta* **1954**, *14*, 574−575.

(60) Berger, A.; Linderstrøm-Lang, K. Deuterium exchange of poly-DL-alanine in aqueous solution. *Arch. Biochem. Biophys.* **1957**, *69*, 106−118.

(61) Privalov, P. L.; Tsalkova, T. N. Micro- and macro-stabilities of globular proteins. *Nature* **1979**, *280*, 693−696.

(62) Englander, S. W.; Mayne, L.; Bai, Y.; Sosnick, T. R. Hydrogen exchange: the modern legacy of Linderstrøm-Lang. *Protein Sci.* **1997**, *6*, 1101−1109.

(63) Huyghues-Despointes, B.; Scholtz, J.; Pace, C. Protein conformational stabilities can be determined from hydrogen exchange rates. *Nat. Struct. Mol. Biol.* **1999**, *6*, 910−912.

(64) Craig, P. O.; Lätzer, J.; Weinkam, P.; Hoffman, R. M. B.; Ferreiro, D. U.; Komives, E. A.; Wolynes, P. G. Prediction of Native-State Hydrogen Exchange from Perfectly Funneled Energy Landscapes. *J. Am. Chem. Soc.* **2011**, *133*, 17463−17472.

(65) Balasubramaniam, D.; Komives, E. A. Hydrogen-exchange mass spectrometry for the study of intrinsic disorder in proteins. *Biochim. Biophys. Acta* **2013**, *1834*, 1202−1209.

(66) Persson, F.; Halle, B. How amide hydrogens exchange in native proteins. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112*, 10383−10388.

(67) Shirley, B. A.; Stanssens, P.; Hahn, U.; Pace, C. N. Contribution of hydrogen bonding to the conformational stability of ribonuclease T1. *Biochemistry* **1992**, *31*, 725−732.

(68) Bahar, I.; Wallqvist, A.; Covell, D. G.; Jernigan, R. L. Correlation between native-state hydrogen exchange and cooperative residue fluctuation from a simple model. *Biochemistry* **1998**, *37*, 1067−1075.

(69) Miller, D. W.; Dill, K. A. A statistical mechanical model for hydrogen exchange in globular proteins. *Protein Sci.* **1995**, *4*, 1860−1873.

(70) Vendruscolo, M.; Paci, E.; Dobson, C. M.; Karplus, M. Rare Fluctuations of Native Proteins Sampled by Equilibrium Hydrogen Exchange. *J. Am. Chem. Soc.* **2003**, *125*, 15686−15687.

(71) Wagner, A. Robustness, evolvability, and neutrality. *FEBS Lett.* **2005**, *579*, 1772−1778.

(72) Arnold, F. H. How proteins adapt: lessons from directed evolution. *Cold Spring Harbor Symp. Quant. Biol.* **2009**, *74*, 41−46.

(73) Maynard Smith, J. Natural Selection and the concept of a protein space. *Nature* **1970**, *225*, 563−564.

(74) Socha, R. D.; Tokuriki, N. Modulating protein stability - directed evolution strategies for improved protein function. *FEBS J.* **2013**, *280*, 5582−5595.

(75) Kaltenbach, M.; Jackson, C. J.; Campbell, E. C.; Hollfelder, F.; Tokuriki, N. Reverse evolution leads to genotypic incompatibility despite functional and active site convergence. *Elife* **2015**, *4*, No. e06492.

(76) Ogbunugafor, C. B. A Reflection on 50 Years of John Maynard Smith's "Protein Space.". *Genetics* **2020**, *214*, 749−754.

(77) Weinreich, D. M.; Delaney, N. F.; DePristo, M. A.; Hartl, D. L. Darwinian evolution can follow only very few mutational paths to fitter proteins. *Science* **2006**, *312*, 111−114.

(78) Sailer, Z. R.; Harms, M. J. Molecular ensembles make evolution unpredictable. *Proc. Natl. Acad. Sci. U. S. A.* **2017**, *114*, 11938−11943.

(79) Kimura, M. Evolutionary rate at the molecular level. *Nature* **1968**, *217*, 624−626.

(80) Bloom, J. D.; Arnold, F. H. In the light of directed evolution: pathways of adaptive protein evolution. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 9995−10000.

(81) Draghi, J. A.; Parsons, T. L.; Wagner, G. P.; Plotkin, J. B. Mutational robustness can facilitate adaptation. *Nature* **2010**, *463*, 353−355.

(82) Kurahashi, R.; Sano, S.; Takano, K. Protein Evolution is Potentially Governed by Protein Stability: Directed Evolution of an Esterase from the Hyperthermophilic Archaeon Sulfolobus tokodaii. *J. Mol. Evol.* **2018**, *86*, 283−292.

(83) Starr, T. N.; Thornton, J. W. Epistasis in protein evolution. *Protein Sci.* **2016**, *25*, 1204−1218.