PNAS COMMENTARY

Evolutionary entrenchment in immune proteins selects against stabilizing mutations

Faruck Morcos[a,b,1]

Characterizing the evolutionary mechanisms responsible for phenotypic traits and functional outcomes is a challenging task, as evolutionary processes act within a complex and dynamic network of interactions leading to extant organisms and natural systems. The case of molecular evolution is no exception, as molecular processes are also part of a web of intricate causal relationships. For the case of proteins, however, the situation is more promising. This is due to our ability to measure protein biophysical properties like folding kinetics and energetics (1, 2) and a mature understanding of the theory of protein folding (3, 4). Additionally, by experimentally probing the enormous, but discrete, mutational space, information about stability and function may be revealed, hence allowing concrete avenues for scientific exploration of evolutionary processes when combining experimental findings with evolutionary data for homologous systems. Taken together, all these factors constitute a framework to tackle the problem of evolutionary characterization in a more systematic way. The implications of achieving a more accurate description of the effects of mutations and its role in evolutionary history are significant, as many protein systems are involved in fundamental processes of life or have direct implications for health and disease.

One example of this is protein S100A9, which in humans contributes to innate immunity by activating Toll-like receptor 4 (TLR4), which, in turn, associates with NF-κB and other pathways to promote proinflammatory activity (5, 6). Although a complete description of all functional mechanisms of S100A9 has not been determined, this calcium-binding protein is thought to use a hydrophobic surface region to interact with TLR4 and modulate proinflammatory activity (5, 7). In a recent study in PNAS, Harman et al. (8) investigate the evolution of this family of proteins and, in particular, how the human homolog (hS100A9) is affected by a single site mutation that provides stability upon calcium binding. Intriguingly, this mutation leads to a loss of proinflammatory activity. This mutation was first identified when the same group used ancestral sequence reconstruction (ASR) to propose candidate ancestors of hS100A9 (9). They found that a mutation at the sequence position 63 from its original methionine (M) to a phenylalanine (F), (M63F hereafter), renders the protein nonfunctional even though phenylalanines has existed at that site in ancient homologs of this protein. The team set to identify the mechanistic causes that make this mutation disruptive, albeit inducing stability. Fig. 1 provides an overview of the biophysical and phenotypical observations that Harman et al. uncovered upon M63F mutation on S100A9 and the experimental and computational framework utilized to provide evidence for these observations. One of their initial observations was that M63F has a negligible effect on the protein energetics when S100A9 is in its *apo* state. The situation changes drastically when larger concentrations of calcium are used. By using unfolding energy measurements and equilibrium unfolding with circular dichroism (CD), they found that M63F is indeed significantly more stable than the wild-type sequence. The mutant required a higher concentration of urea to unfold, possessed slower unfolding kinetics, and was more resistant to proteolysis. Given that the cellular environment is calcium-rich, they reasoned that such acquired stability could, in fact, be disruptive for proinflammatory activity, and more concretely, it might affect the interaction with TLR4. To provide support for this hypothesis, they performed hydrogen–deuterium exchange (HDX) experiments and 3D nuclear magnetic resonance (NMR) to identify dynamic 3D coordinates of the mutant and analyze how hydrophobic surfaces in the protein changed upon mutation. They unveiled how this mutation has a structural effect on the organization of helices 3 and 4 of the protein (see H3 and H4 shown in Fig. 1).

This reorganization was manifested in the compaction of a hydrophobic patch that presumably interacts with TLR4. M63F induces a reduction in the angle formed between H3 and H4. The angle decreased from 68° to 38°, leading to a smaller exposed hydrophobic region. In order to reveal the causes of such structural and biochemical change, they turned to residue–residue interactions near position 63. They realized that residue 37, which is also a phenylalanine, formed a tighter interaction with position 63, potentially forming an edge-face π-stack. Using this as their leading hypothesis for stabilization, they decided to do a series of mutational experiments to characterize the 63–37 pair. They found that mutating position 37 to leucine abolishes the stabilization introduced in M63F. They provided conclusive support to this claim by doing HDX and chemical denaturation experiments (see Fig. 1). Having established that positions 63 and 37 form a "molecular staple," Harman et al. questioned whether such stabilizing interaction leads, in fact, to a loss of functional activity. The group did this by probing NF-κB activation

Author affiliations: [a]Department of Biological Sciences, University of Texas at Dallas, Richardson, TX 75080; and [b]Center for Systems Biology, University of Texas at Dallas, Richardson, TX 75080

Author contributions: F.M. wrote the paper.

The author declares no competing interest.

See companion article, "Evolution avoids a pathological stabilizing interaction in the immune protein S100A9," 10.1073/pnas.2208029119.

[1]Email: faruckm@utdallas.edu.

This article contains supporting information online at http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2216087119/-/DCSupplemental.

Published November 30, 2022.

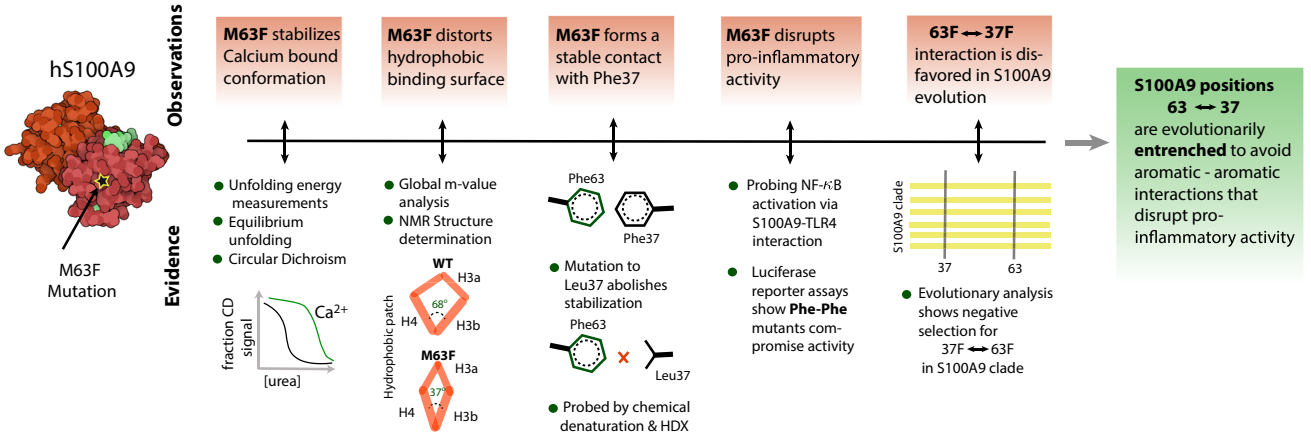

Fig. 1. The S100A9 M63F mutant leads to stabilization and loss of function. A mutation at position 63 in hS100A9 stabilizes a calcium-bound conformation. This stabilization is due to a "molecular staple" that arises between positions 63 and 37. This interaction leads to a reduction of a hydrophobic patch that disrupts association with TLR4 and ultimately abolishes proinflammatory activation. Evolutionary analysis reveals that this interaction has been negatively selected in the S100A9 clade, presumably due to its deleterious effect in function.

via the S100A9–TLR4 interaction using a luciferase reporter assay. They concluded that, introducing aromatic–aromatic interactions leads to a disruption in proinflammatory activity.

It is clear that this collection of observations and their corresponding experimental support provides a rigorous and detailed biophysical picture of the deleterious effect of this stabilizing mutation. However, it still left open the question whether effects could also be present in members of the S100A9 clade or in any of their ancient relatives. To tackle this question, Harman et al. analyzed the statistics of the members of the S100A9 clade as well as the sequences of other members of the family. Although they found that phenylalanines at positions 63 and 37 have been observed, they are statistically underrepresented as a pair in the S100A9 clade. Given that this phenomenon is not observed in other homologs, e.g., S100A8, S100A12, or MRP-126, they concluded that the acquired role of S100A9 in innate immunity constrained these two positions to negatively select this pair. This is one example of a phenomenon called entrenchment in the field of molecular evolution (10), whereby a site which allowed a particular amino acid in a given position in the past is now constrained to avoid such mutation once a novel functional constraint has appeared. In this case the stabilization provided by M63F in the *holo* state of the protein has been avoided for the S100A9 clade as it disrupts its role in proinflammatory activities. For the case of ancient homologs, this restriction is not present, and the statistics of simultaneous aromatic residues at positions 63 and 37 are only driven by drift. As these effects cannot be captured by analyzing the site statistics in an independent way, these observations provide further evidence of the importance of epistasis in protein evolution and function (11).

The work of Harman et al. provides a rare view of the intricacies of molecular evolution. The availability of powerful experimental tools like CD, HDX, and 3D NMR as well as the access to evolutionary data and analysis techniques has boosted the potential to clearly characterize the effects of specific mutations in an evolutionary context. Research in providing concrete experimental cases at evolutionary phenomena like entrenchment and evolvability (12), among several others in the field, includes recent work by Ding et al. (13) in toxin/antitoxin systems. Moreover, theoretical work that learns from extant sequences provides insights into context dependence in evolution, and epistasis (11, 14) shows promise in modeling these phenomena in a generalized way. Although much further work is still needed, disentangling evolutionary dilemmas, e.g., evolution against stability, using the lens of biophysics has proven a promising approach (15). This type of work supports our view of evolution as a multilayered phenomenon, where a large number of constraints on different dynamic structures, their partners, and their context must be accounted for to disentangle the mechanisms of function. As Dobzhansky stated, "Nothing in biology makes sense except in the light of evolution" (16).

ACKNOWLEDGMENTS. My research in protein evolution is supported by grants from the National Institute for General Medical Sciences of the NIH (grant R35GM133631) and the Division of Molecular and Cellular Biosciences (MCB) of the NSF (grant MCB-1943442).

1. V. Munoz, W. A. Eaton, A simple model for calculating the kinetics of protein folding from three-dimensional structures. *Proc. Natl. Acad. Sci. U.S.A.* **96**, 11311–11316 (1999).
2. S. Marqusee, V. H. Robbins, R. L. Baldwin, Unusually stable helix formation in short alanine-based peptides. *Proc. Natl. Acad. Sci. U.S.A.* **86**, 5286–5290 (1989).
3. J. N. Onuchic, Z. Luthey-Schulten, P. G. Wolynes, Theory of protein folding: The energy landscape perspective. *Annu. Rev. Phys. Chem.* **48**, 545–600 (1997).
4. K. A. Dill, S. B. Ozkan, M. S. Shell, T. R. Weikl, The protein folding problem. *Annu. Rev. Biophys.* **37**, 289–316 (2008).
5. Z. He *et al.*, Cd14 is a co-receptor for tlr4 in the s100a9-induced pro-inflammatory response in monocytes. *PLoS One* **11**, e0156377. (2016).
6. H. Gao *et al.*, S100a9-induced release of interleukin (il)-6 and il-8 through toll-like receptor 4 (tlr4) in human periodontal ligament cells. *Mol. Immunol.* **67**, 223–232 (2015).
7. A. N. Loes, J. T. Bridgham, M. J. Harms, Coevolution of the toll-like receptor 4 complex with calgranulins and lipopolysaccharide. *Front. Immunol.* **9**, 304 (2018).
8. J. L. Harman *et al.*, Evolution avoids a pathological stabilizing interaction in the immune protein s100a9. *Proc. Natl. Acad. Sci. U.S.A.*, (2022).
9. J. L. Harman *et al.*, Evolution of multifunctionality through a pleiotropic substitution in the innate immune protein s100a9. *Elife* **9**, (2020).
10. P. Shah, D. M. McCandlish, J. B. Plotkin, Contingency and entrenchment in protein evolution under purifying selection. *Proc. Natl. Acad. Sci. U.S.A.* **112**, E3226–E3235 (2015).
11. J. A. de la Paz, C. M. Nartey, M. Yuvaraj, F. Morcos, Epistatic contributions promote the unification of incompatible models of neutral molecular evolution. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 5873–5882 (2020).
12. M. Kirschner, *J. Gerhart. Evolvability. Proc. Natl. Acad. Sci. U.S.A.* **95**, 8420–8427 (1998).
13. D. Ding *et al.*, Co-evolution of interacting proteins through non-contacting and non-specific mutations. *Nat. Ecol. Evol.* **6**, 590–603 (2022).

14. M. Bisardi, J. Rodriguez-Rivas, F. Zamponi, M. Weigt, Modeling sequence-space exploration and emergence of epistatic signals in protein evolution. *Mol. Biol. Evol.* **39** (2022).
15. T. Sikosek, H. S. Chan, Biophysics of protein evolution and evolutionary protein biophysics. *J. R. Soc. Interface* **11**, 20140419 (2014).
16. T. Dobzhansky, Nothing in biology makes sense except in the light of evolution. *Am. Biol. Teach.* **35**, 125–129 (1973).