The Evolution of Minimal Specificity

Do protein interactions evolve towards minimal specificity?

To answer this question, I simulated boolean logic circuits to model binding interactions between proteins and ligands. I then applied a genetic algorithm to examine the evolutionary behavior of binding specificity. My results are consistent with the hypothesis that minimal specifity is an emergent property of evolution, but I also observed that minimal specificity can be thwarted when promiscuous ligands appear in the environment.

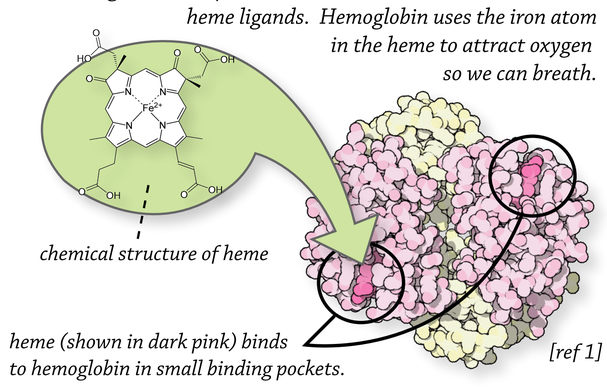
by Victor Hanson-Smith

Department of Computer and Information Science, University of Oregon
Center for Ecology and Evolutionary Biology, University of Oregon
Contact: victorhs@cs.uoregon.edu

Evolutionary biologists often observe that living systems evolve "only as much as they need to" in response to environmental changes. For example, some proteins bind ligands — i.e., small organic molecules — using ligand-binding surfaces with a minimal number of chemical bonds to disambiguate particular target ligands from other environmental ligands. In other words, some proteins are minimally specific for their target ligands. ***Is minimal specificity an emergent property of evolution? If so, under what conditions do these minimal interactions evolve?***

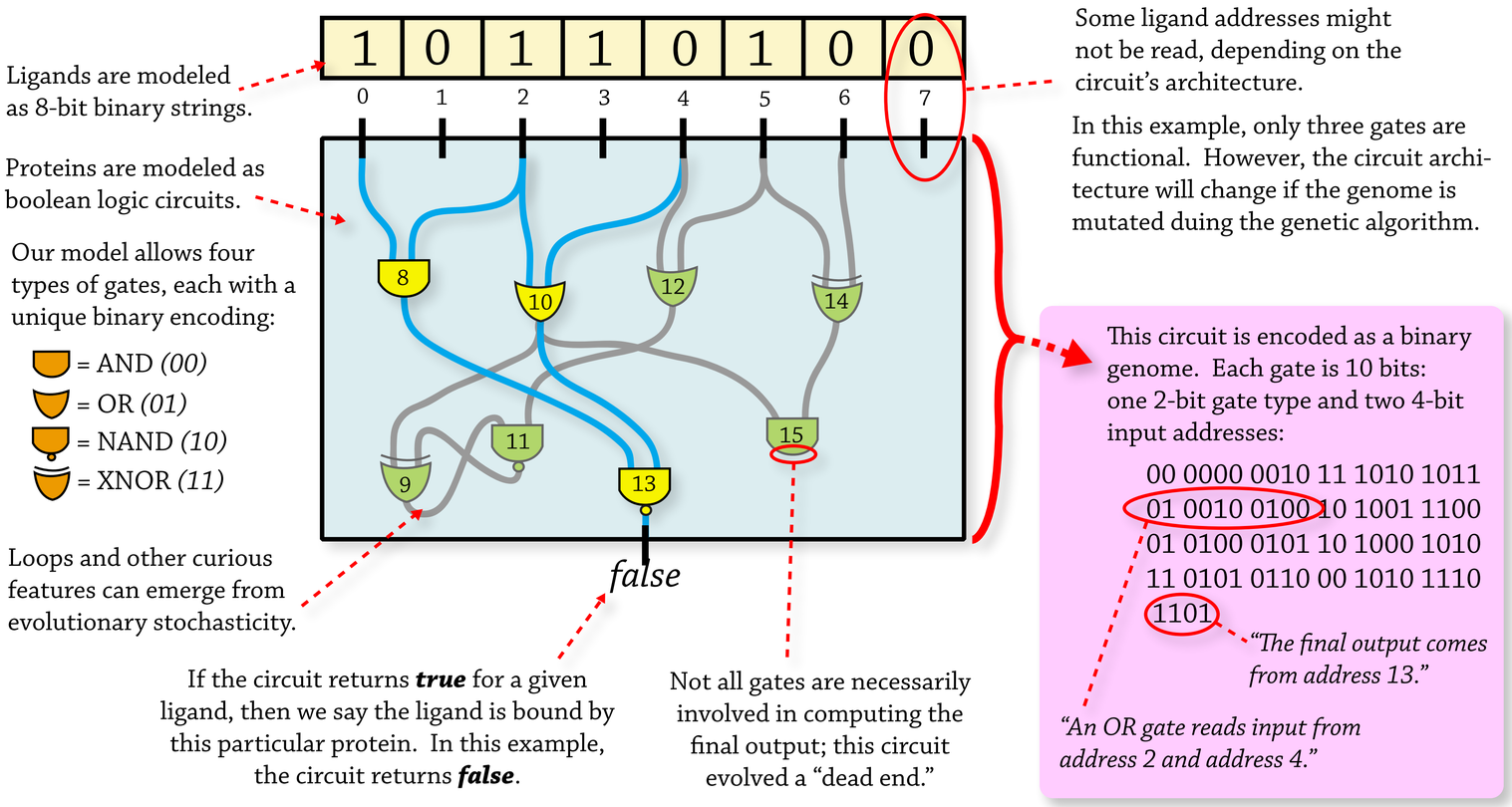
Ligand-binding, an example:

Heme is a ligand, circulating in your bloodstream. Hemoglobin is a protein that forms a covalent bond with heme ligands. Hemoglobin uses the iron atom in the heme to attract oxygen so we can breath.



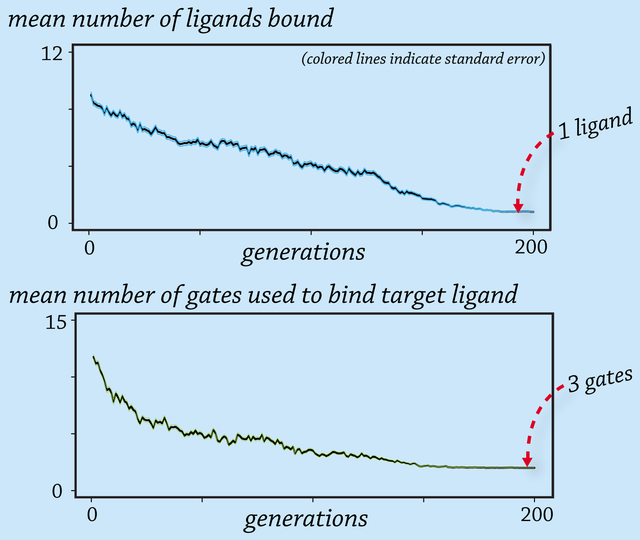
The Model

Protein behavior is notoriously challenging to simulate because the number of biophysical interactions within a single protein often exceed the limits of computational tractability. As an alternative, I modeled proteins as boolean logic circuits and I modeled ligands as binary strings. If a "protein" returns the value ***true*** for a given "ligand," then we say this protein positively binds this ligand. Furthermore, each protein can be encoded in a binary genome, and a population of genomes can be evolved using a genetic algorithm to select those individuals with fitness for binding a particular target ligand. This model is inspired by previous work from Uri Alon's lab [*refs 2,3,4,5*].

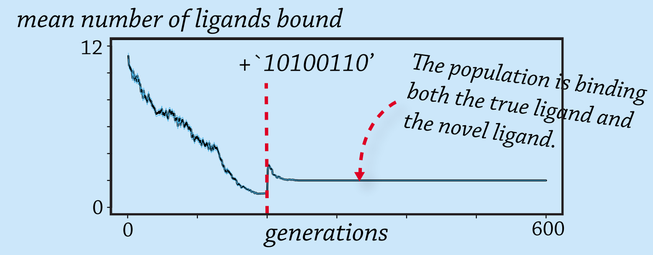


I used a genetic algorithm to evolve a population of 1000 proteins to bind the target ligand `10110110' against 19 other non-target ligands. The fitness of a single protein was meausred as the inverse proportion of non-target ligands bound. I used a crossover recombination strategy and also randomly mutated 10% of genome sites at each generation.

I observed that a population of simulated proteins can indeed evolve to find a minimally specific binding solution. Here the population converged on one of several possible three-gate solutions; for this problem, three gates is the minimum required to disambiguate the target ligand from the other non-target ligands.



I then introduced a novel ligand (`10100110') that was promiscuously bound by the population at generation 200. At this point, the population was no longer minimally specific because the evolved logic circuitry could not disambiguate between the novel ligand and the true target ligand. Furthermore, the population was effectively trapped on a local fitness minima; to find a new minimally specific solution, the population must accumulate a series of extremely low probability neutral mutations.



This scenario demonstrates why some living systems struggle to evolutionarily adapt to radical biochemical environmental changes, such as the introduction of man-made exogenous chemicals.

Minimal Specificity

The idea of **specificity** was introduced in the artificial intelligence community during research into fuzzy logic [*refs 6,7*]. Specificity measures the degree to which a fuzzy set points to only one element as its member.
A fuzzy set is **minimally specific** [*refs 8,9*] if and only if:
(1) the set points to only one element
(2) the set uses a minimum number of logic rules to positively identify the single element.

My key insight is that ligand-binding surfaces demonstrate fuzzy logic to discriminate between ligands. A surface might partially-bind many different ligands, but fully-binds only a few particular ligands. A ligand-binding surface is said to be specific if it binds only one target ligand. Inversely, a ligand-binding surface is **promiscuous** if it positively binds several ligands.

In terms of our model, a simulated "protein" is minimally specific if (1) it returns true for the target ligand and false for all other ligands, and (2) the protein's internal circuit architecture computes the *true/false* outcome using a minimum number of logic gates.

References

[1] (image from...) Bolton, W., Perutz, M.F. (1970) Three dimensional fourier synthesis of horse deoxyhaemoglobin at 2.8 Angstrom units resolution. Nature 228: 551-552 doi:10.2210/pdb2dhb/pdb
[2] "Spontaneous Evolution of Modularity and Network Motifs." Kashtan and Alon., Proceedings of the National Academy of Science, 2005.
[3] "Varying Environments can Speed Up Evolution." Kashtan et al., Proceedings of the National Academy of Science, 2007.
[4] "Facilitated Variation: How Evolution Learns from Past Environments to Generalize to New Envrionments." Parter et al., PLoS Computational Biology, 2008
[5] "Extinctions in Heterogenous Environments and the Evolution of Modularity." Kashtan et al., Evolution, 2009.
[6] "A Mathematical Theory of Evidence." Shafer, Princeton University Press, 1976.
[7] "Entropy and Specificity in a Mathemtical Theory of Evidence." R.R. Yager, International Journal of General Systems, 1983, 9:249-260
[8] "The Principle of Minimum Specificity as a Basis for Evidential Reasoning." Dubois and Prade, Uncertainty in Knowledge-Based Systems, Springer-Verlag, 1986.
[9] "A Set-Theoretic View of Belief Functions." Dubois and Prade, Studies in Fuzziness and Soft-Computing. Springer-Verlag, 1986.