Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Biotechnology

There's (still) plenty of room at the bottom

Noah Olsman[1] and Lea Goentoro[2]

Motifs, circuits, and networks are core conceptual elements in modern systems and synthetic biology. While there are still undoubtedly more fascinating computations to discover at network level, there are also rich computations that we are only beginning to uncover within the diverse molecules that constitute the networks. Here we explore some work, both new and old, that showcases the incredible computational capacity of seemingly simple molecular mechanisms. A more sophisticated understanding of computations at the molecular level will inspire the development of a more nuanced toolbox for future biological engineering.

Addresses
[1] Department of Computing and Mathematical Sciences, California Institute of Technology, United States
[2] Division of Biology and Biological Engineering, California Institute of Technology, United States

Corresponding authors: Olsman, Noah (nolsman@caltech.edu), Goentoro, Lea (goentoro@caltech.edu)

Current Opinion in Structural Biology 2018, **54**:72–79

This review comes from a themed issue on **Analytical biotechnology**

Edited by **Hiroshi Shimizu** and **Fumio Matsuda**

https://doi.org/10.1016/j.copbio.2018.01.029

Exploring the molecular toolbox

Almost 60 years ago, Richard Feynman gave a now-famous lecture titled 'There's Plenty of Room at the Bottom' [1•], where he argued that we had only just begun to understand the extent to which the physical world can be manipulated at the molecular scale. He expressed wonderment at processes like photosynthesis and the translation of proteins. At the time, relatively little was known about the structure, function, and organization of the molecules that underly these phenomena, but it was clear from their design that evolution is a resourceful engineer (Figure 1).

In the last few decades, Feynman's vision has largely become a reality. We now understand, in a much deeper way, that single cells have an astonishing capacity to sense and make decisions about their environment. From quorum sensing in bacteria to embryonic development in animals, a great deal of information must be processed using DNA, RNA, and proteins. When studying information processing in biology, we often focus on the computational capacity of circuits and motifs of a few components [2] and networks of tens to hundreds of elements [3–5]. While there is certainly still much to be explored at the level of circuits and networks, this perspective often coarse-grains the finer molecular details of these systems.

Our viewpoint here will zoom in on these often-overlooked molecular details, and highlight several case studies where molecules perform an impressive range of computation. For instance, while a membrane receptor can simply be viewed as serving to transduce signal from outside the cell, a closer examination reveals that it can be an adaptive component in signaling, performing a wide range of nonlinear computation. Similarly, mRNA is a template for protein translation, but a closer look reveals it to be a powerful regulatory hub, integrating a variety of chemical and environmental stimuli.

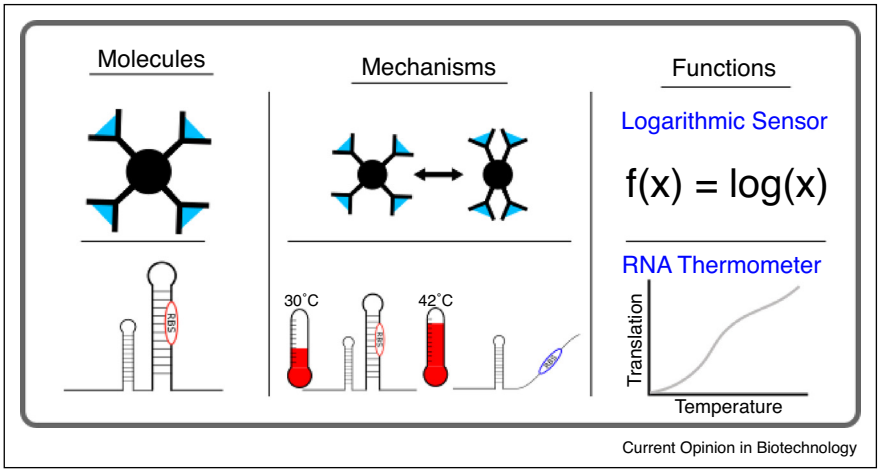
An enzymatic proofreader

We begin with a seminal paper by John Hopfield, published in 1973 [6], that asked a simple question: how is it that the synthesis of biological molecules has such a low rate of error? Hopfield considered a simple model of protein translation. Protein translation proceeds by an mRNA template being processed by a ribosome, which recruits tRNA that ultimately attach amino acids to the nascent chain. A key step is when a ligase attaches an amino acid to a transfer RNA (tRNA). If the ligase attaches the wrong amino acid then, even if the correct tRNA recruited, there will be an error in protein translation.

Hopfield first analyzed the expected error rate due to the ligation of incorrect amino acids to tRNA. Take, for example, equal concentrations of the amino acids isoleucine and valine. The binding affinity of the isoleucine ligase for the former is about 100× that of the latter, implying an approximate error rate of 10^{-2}. With more types of amino acids, one would expect the error rate to only get worse. These results are in stark contrast to the real error rate in protein translation, which is closer to 10^{-4} [7].

The key to resolving this discrepancy is the phenomenon that Hopfield referred to as kinetic proofreading. To motivate this idea, imagine a point during the assembly of a protein where amino acid • is required, however there is an equal amount of amino acid • available, described by the reactions in Figure 2a. Let us assume that the

Figure 1



Molecules, mechanisms, and functions. A primary goal of systems and synthetic biology is to gain a deep understanding of the connection between the molecules that make up biological processes and their associated functions. While we often focus on large biomolecular networks, we can often gain insight from studying the quantitative functional properties of individual molecular mechanisms. In the row, we see a protein with multiple subunits, the mechanism of conformational switching, and the function of responding to signal in a logarithmic fashion. On the bottom row, we see an RNA with a particular secondary structure, the mechanism of temperature regulating the availability of the ribosome binding site, and the function of protein translation being a function of temperature.

ligase $L_{\bullet}$ has a relative dissociation rate $\frac{K_{\bullet}}{K_{\bullet}} = 100$. This would yield the 10^{-2} error rate mentioned earlier.

Now imagine adding a second step that we will refer to as *thermodynamic* proofreading. Instead of directly adding an amino acid, we will instead have an intermediary stage where some form of post-translational modification takes place. Ideally this intermediate stage would allow for additional ligand specificity, giving the ligase an extra opportunity to ameliorate erroneous binding (Figure 2b). Unfortunately, at equilibrium this cannot be the case. Hopfield showed explicitly that due to the constraints of thermodynamic equilibrium, while the second step does preferentially reverse off target binding, the nature of the first step implies that there will be far more $L_{\bullet}-\bullet$ complexes than $L_{\bullet}-\bullet$ complexes. These rates *must* balance in such a way that the rate of erroneous binding is never less than the original 10^{-2}.

Since the impediment comes from thermodynamic constraints, it makes sense to ask if we can somehow sidestep them. One way to do this is to pump energy into the system, by some ATP-consuming enzymatic process, which means we are in the regime of chemical *kinetics* (Figure 2c). This frees us of thermodynamic equilibrium constraints and makes it possible to reduce error rates by a factor of the previous rate squared, or 10^{-4}. The addition of more successive energy-intensive steps would continue to drop the error rate. Over the last 40 years kinetic proofreading has emerged as a pervasive mechanism in various biological pathways, from T-cell receptor signal transduction [8] to chromatin remodeling [9,10]. On the theoretical side, recent work has vastly expanded our understanding of how kinetic proofreading works from the perspective of non-equilibrium statistical mechanics [11•]. While there is no free lunch in life, sometimes it is worth paying a bit extra to get a better meal.

An RNA thermometer

Next we shift focus from preventing errors to preventing catastrophes. Every known organism on Earth has some mechanism for responding to a sudden increase in temperature, referred to as heat shock response (HSR) [12]. One reason heat shock is dangerous is that it can cause proteins to become misfolded, which can be lethal for a cell. To avoid this, the HSR system senses temperature change, and synthesizes chaperone proteins whose job is it is to refold proteins [13].

How does a cell sense when it needs to make chaperones? One way is to sense misfolded proteins (i.e. feedback response), the other is to sense temperature changes so that chaperone synthesis can begin before misfolded proteins have already accumulated (i.e. feedforward response). While both sensing mechanisms exist [14,15], we focus here on the latter feedforward response that is mediated by messenger RNA.

Since heat shock response needs to be fast, it is reasonable to posit that any temperature sensing mechanism must be directly linked to the synthesis of chaperones. To solve this, *Escherichia coli* have evolved an ingenious