Statistical physics of self-replication

Jeremy L. England

Articles you may be interested in

Perspective: Reaches of chemical physics in biology
The Journal of Chemical Physics **139**, 121701121701 (2013); 10.1063/1.4820139

Statistical physics of self-replication

Jeremy L. England

Department of Physics, Massachusetts Institute of Technology, Building 6C, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139, USA

Self-replication is a capacity common to every species of living thing, and simple physical intuition dictates that such a process must invariably be fueled by the production of entropy. Here, we undertake to make this intuition rigorous and quantitative by deriving a lower bound for the amount of heat that is produced during a process of self-replication in a system coupled to a thermal bath. We find that the minimum value for the physically allowed rate of heat production is determined by the growth rate, internal entropy, and durability of the replicator, and we discuss the implications of this finding for bacterial cell division, as well as for the pre-biotic emergence of self-replicating nucleic acids. © 2013 Author(s). All article content, except where otherwise noted, is licensed under a Creative Commons Attribution 3.0 Unported License. [http://dx.doi.org/10.1063/1.4818538]

INTRODUCTION

Every species of living thing can make a copy of itself by exchanging energy and matter with its surroundings. One feature common to all such examples of spontaneous "self-replication" is their statistical irreversibility: clearly, it is much more likely that one bacterium should turn into two than that two should somehow spontaneously revert back into one. From the standpoint of physics, this observation contains an intriguing hint of how the properties of self-replicators must be constrained by thermodynamic laws, which dictate that irreversibility is always accompanied by an increase of entropy. Nevertheless, it has long been considered challenging to speak in universal terms about the statistical physics of living systems because they invariably operate very far from thermodynamic equilibrium, and therefore need not obey a simple Boltzmann probability distribution over microscopic arrangements. Faced with such unconstrained diversity of organization, it is quite reasonable to worry that the particular mechanistic complexity of each given process of biological self-replication might overwhelm our ability to say much in terms of a general theory.

For the beginnings of a way forward, we should consider a system of fixed particle number N and volume V in contact with a heat bath of inverse temperature β. If we give labels to the microstates of this system i, j, etc. and associate energies E_i, E_j, etc., respectively with each such microstate, then the underlying time-reversal symmetry of Hamiltonian dynamics tells us that the following detailed balance relation holds at thermal equilibrium:[1]

$$\frac{e^{-\beta E_i}}{Z(\beta)}\pi(i \rightarrow j; \tau) = \frac{e^{-\beta E_j}}{Z(\beta)}\pi(j^* \rightarrow i^*; \tau). \qquad (1)$$

Here, i^* is the momentum-reversed partner of i, and $Z(\beta)$ is the canonical partition function of the system. The transition matrix element $\pi(i \rightarrow j; \tau)$ is the conditional probability that the system is found to be in microstate j at time $t = \tau > 0$ given that it started off in microstate i at an earlier time

$t = 0$. Thus, the above relation states that when the system has relaxed to thermal equilibrium and achieved a Boltzmann distribution over microstates ($p(i) = \frac{e^{-\beta E_i}}{Z(\beta)}$), the probability currents connecting i to j in the forward and time-reversed directions are equal.

Progress comes from recognizing that the heat expelled into the bath over a transition from i to j is given by $\Delta Q_{i \rightarrow j} = E_i - E_j$, and moreover that the quantity $\beta \Delta Q_{i \rightarrow j}$ is the amount by which the entropy of the heat bath changes over the course of this transition. Thus, we may write

$$\frac{\pi(j^* \rightarrow i^*; \tau)}{\pi(i \rightarrow j; \tau)} = \exp\left[-\Delta S_{bath}^{i \rightarrow j}\right]. \qquad (2)$$

Equation (2) sets up a microscopically detailed relationship between the irreversibility of a transition (that is, how much more likely it is to happen in the forward direction than in the reverse direction) and the amount entropy is increased in the surroundings over the course of the forward trajectory. Moreover, it should be stressed that while this result is derived from a statement of detailed balance (which only holds at equilibrium) it is itself valid for any transition between two microstates, and thus applies to the relaxation dynamics of undriven systems arbitrarily far from equilibrium.

The principal aim of this work is to show that the microscopically detailed, quantitative relationship between irreversibility and entropy production illustrated above has significant, general thermodynamic consequences for far-from-equilibrium, macroscopic processes such as biological self-replication. Building on past results in nonequilibrium statistical mechanics,[2] we will first derive a generalization of the Second Law of Thermodynamics for macroscopic irreversible processes. Subsequently, we will use this result to demonstrate a lower bound on the heat output of a self-replicator in terms of its size, growth rate, internal entropy, and durability. We will furthermore show, through analysis of empirical data, that this bound operates on a scale relevant to the functioning of real microorganisms and other self-replicators.

MACROSCOPIC IRREVERSIBILITY

Our goal is to determine thermodynamic constraints on a macroscopic transition between two arbitrarily complex coarse-grained states. To do so, we have to think first about how probability, heat, and entropy are related at the microscopic level. Equation (2) established such a relationship in one special case, but the linkage turns out to be far more general, for it has been shown[2] that for all heat bath-coupled, time-symmetrically driven nonequilibrium systems whose dynamics are dominated by diffusive motions that lack any sense of ballistic inertia, any trajectory starting at $x(0)$, and going through microstates $x(t)$ over time τ obeys

$$\beta \Delta Q[x(t)] = \ln \left[\frac{\pi[x(t)]}{\pi[x(\tau - t)]} \right]. \tag{3}$$

Here, $\beta \equiv 1/T$ sets the temperature T of the bath in natural units, $\pi[x(t)]$ is the probability of the trajectory $x(t)$ given $x(0)$, and ΔQ is the heat released into the bath over the course of $x(t)$. This formula captures the essential microscopic relationship between heat and irreversibility: the more likely a forward trajectory is than its time-reverse, the more the entropy of the universe is increased through the exhaustion of heat into the surrounding bath. Furthermore, it should be emphasized that this result holds for systems subject to time-symmetric external driving fields. This means that the steady-state probability distribution of the system need not be the Boltzmann distribution, nor even a distribution whose steady-state probability flux satisfies detailed balance, in order for this fundamental relationship between irreversibility and entropy production to hold.

By fixing the starting and ending points of our trajectory ($x(0) = i$, $x(\tau) = j$), we can average the exponential weight of the forward heat over all paths from i to j and obtain $\pi(j \to i; \tau)/\pi(i \to j; \tau) = \langle \exp[-\beta \Delta Q^{\tau}_{i \to j}] \rangle_{i \to j}$. This microscopic rule (of which Eq. (2) is clearly the special, undriven case) must have macroscopic consequences, and to investigate these we have to formalize what it means to talk about our system of interest in macroscopic terms. In order to do so, we first suppose there is some coarse-grained, observable condition **I** in the system, such as the criterion "The system contains precisely one healthy, exponential-growth-phase bacterium at the start of a round of cell division." If we prepare the system under some set of controlled conditions and then find that the criterion for **I** is satisfied, we can immediately associate with **I** a probability distribution $p(i|\mathbf{I})$ which is the implicit probability that the system is in some particular microstate i, given that it was prepared under controlled conditions and then observed to be in macrostate **I**.

Suppose now that we let some time interval τ pass while keeping our system in contact with a heat bath of inverse temperature β (In subsequent expressions for $\pi(i \to j)$ the implicit τ will be omitted). At that point, we could introduce a second criterion for coarse-grained observation of the system (e.g., "The system contains precisely *two* healthy, exponential-growth-phase bacteria at the start of cell division."), which we might call **II**, and in the event that we observed this criterion to be satisfied, we could immediately define the probability distribution $p(j|\mathbf{II})$ as being the likelihood of the system being in microstate j given that a system initially prepared in **I** subsequently propagated over a certain period of time τ and was then found to be in macrostate **II**.

Having constructed a probabilistic definition of our macrostates of interest, we may now also define associated quantities that will allow us to give a macroscopic definition to irreversibility. In particular, we can write

$$\pi(\mathbf{I} \to \mathbf{II}) = \int_{\mathbf{II}} dj \int_{\mathbf{I}} di \, p(i|\mathbf{I}) \pi(i \to j), \tag{4}$$

and

$$\pi(\mathbf{II} \to \mathbf{I}) = \int_{\mathbf{I}} di \int_{\mathbf{II}} dj \, p(j|\mathbf{II}) \pi(j \to i). \tag{5}$$

The first of these probabilities $\pi(\mathbf{I} \to \mathbf{II})$ gives the likelihood that a system prepared according to **I** is observed to satisfy **II** after time τ. The second probability $\pi(\mathbf{II} \to \mathbf{I})$ gives the likelihood that after *another* interval of τ that the same system would be observed again to satisfy **I**. Putting these two quantities together and taking their ratio thus quantifies for us the irreversibility of spontaneously propagating from **I** to **II**:

$$
\begin{aligned}
\frac{\pi(\mathbf{II} \to \mathbf{I})}{\pi(\mathbf{I} \to \mathbf{II})} &= \frac{\int_{\mathbf{I}} di \int_{\mathbf{II}} dj \left(\frac{p(j|\mathbf{II})}{p(i|\mathbf{I})} \right) p(i|\mathbf{I}) \pi(j \to i)}{\int_{\mathbf{I}} di \int_{\mathbf{II}} dj \, p(i|\mathbf{I}) \pi(i \to j)} \\
&= \frac{\int_{\mathbf{I}} di \int_{\mathbf{II}} dj \, p(i|\mathbf{I}) \pi(i \to j) \frac{\langle e^{-\beta \Delta Q_{ij}} \rangle_{i \to j}}{e^{\ln\left[\frac{p(i|\mathbf{I})}{p(j|\mathbf{II})}\right]}}}{\int_{\mathbf{I}} di \int_{\mathbf{II}} dj \, p(i|\mathbf{I}) \pi(i \to j)} \\
&= \left\langle \frac{\langle e^{-\beta \Delta Q_{ij}} \rangle_{i \to j}}{e^{\ln\left[\frac{p(i|\mathbf{I})}{p(j|\mathbf{II})}\right]}} \right\rangle_{\mathbf{I} \to \mathbf{II}},
\end{aligned}
\tag{6}
$$

where $\langle \ldots \rangle_{\mathbf{I} \to \mathbf{II}}$ denotes an average over all paths from some i in the initial ensemble **I** to some j in the final ensemble **II**, with each path weighted by its likelihood. Defining the Shannon entropy S for each ensemble in the usual manner ($S \equiv -\sum_i p_i \ln p_i$), we can construct $\Delta S_{int} \equiv S_{\mathbf{II}} - S_{\mathbf{I}}$, which measures the internal entropy change for the forward reaction. Since $e^x \geq 1 + x$ for all x, we may rearrange (6) to write

$$\left\langle e^{-\ln\left[\frac{\pi(\mathbf{II} \to \mathbf{I})}{\pi(\mathbf{I} \to \mathbf{II})}\right] + \ln\left[\frac{p(j|\mathbf{II})}{p(i|\mathbf{I})}\right]} \langle e^{-\beta \Delta Q_{i \to j}} \rangle_{i \to j} \right\rangle_{\mathbf{I} \to \mathbf{II}} = 1 \tag{7}$$

and immediately arrive at

$$\beta \langle \Delta Q \rangle_{\mathbf{I} \to \mathbf{II}} + \ln \left[\frac{\pi(\mathbf{II} \to \mathbf{I})}{\pi(\mathbf{I} \to \mathbf{II})} \right] + \Delta S_{int} \geq 0. \tag{8}$$

If **I** and **II** corresponded to identical groups that each contained all microstates i, we would have $\pi(\mathbf{II} \to \mathbf{I}) = \pi(\mathbf{I} \to \mathbf{II}) = 1$, and the above relation would reduce to a simple statement of the Second Law of Thermodynamics: on average, the change in the entropy of the system ΔS_{int} plus the change in the entropy of the bath $\beta \langle \Delta Q \rangle_{\mathbf{I} \to \mathbf{II}}$ must be greater than or equal to zero, that is, the average total entropy change of the universe must be positive. What we have shown here using Crooks' microscopic relation, however, is that the *macroscopic* irreversibility of a transition from an arbitrary ensemble of states $p(i|\mathbf{I})$ to a future ensemble $p(j|\mathbf{II})$ sets a stricter bound on the associated entropy production: the more irreversible the macroscopic process (i.e., the more negative

$\ln\left[\frac{\pi(\mathbf{II}\rightarrow\mathbf{I})}{\pi(\mathbf{I}\rightarrow\mathbf{II})}\right]$), the more positive must be the minimum total entropy production. Moreover, since the formula was derived under very general assumptions, it applies not only to self-replication but to a wide range of transitions between coarse-grained starting and ending states. In this light, the result in Eq. (8) is closely related to past bounds set on entropy production in information-theoretic terms,[3–5] as well as to the well-known Landauer bound for the heat generated by the erasure of a bit of information.[6]

GENERAL CONSTRAINTS ON SELF-REPLICATION

The generalization of the Second Law presented above applies in a wide range of far-from-equilibrium, heat-bath-coupled scenarios, from the operation of computers to the action of molecular motors. The result turns out to be particularly valuable, however, as a lens through which to take a fresh look at the phenomenon of biological self-replication. Interest in the modeling of evolution long ago gave rise to a rich literature exploring the *consequences* of self-replication for population dynamics and Darwinian competition. In such studies, the idea of Darwinian "fitness" is frequently invoked in comparisons among different self-replicators in a non-interacting population: the replicators that interact with their environment in order to make copies of themselves fastest are more "fit" by definition because successive rounds of exponential growth will ensure that they come to make up an arbitrarily large fraction of the future population.[7]

The first thing to notice here is that exponential growth of the kind just described is a highly irreversible process: in a selective sweep where the fittest replicator comes to dominate in a population, the future almost by definition looks very different from the past. Happily, we are now in the position to be quantitative about the thermodynamic consequences of this irreversibility. Thus, let us suppose there is a simple self-replicator living at inverse temperature β whose population $n \gg 1$ obeys a master equation of the form

$$\dot{p}_n(t) = gn[p_{n-1}(t) - p_n(t)] - \delta n[p_n(t) - p_{n+1}(t)], \quad (9)$$

where $p_n(t)$ is the probability of having a population of n at time t, and $g > \delta > 0$. For simplicity, we assume that a decay event mediated by the rate parameter δ amounts to a reversion of the replicator back into the exact set of reactants in its environment out of which it was made, and we also assume that forming one new replicator out of such a set of reactants changes the internal entropy of the system by an amount Δs_{int}, and on average puts out an amount of heat Δq into the surroundings. We will furthermore ignore in this discussion the additional layer of complexity that comes with considering the effects of spontaneous mutation in evolving populations.[7]

For $n(t = 0) \gg 1$, we expect the behavior of the system to follow a deterministic exponential growth path of the form $n(t) = n(0)e^{(g - \delta)t}$. Whatever the exact value of n, the probability in a short period of time dt that one particular replicator gives birth $\pi(\mathbf{I} \rightarrow \mathbf{II})$ should be $g\,dt$, while the probability of a newly born copy decaying back whence it came in the same-length interval time ($\pi(\mathbf{II} \rightarrow \mathbf{I})$) should be $\delta\,dt$. Thus,

plugging into (8), we have

$$\Delta s_{tot} \equiv \beta\Delta q + \Delta s_{int} \geq \ln[g/\delta]. \quad (10)$$

As an aside, it should be noted here that we have avoided including in the above a spurious multiplicative factor from the number of particles in the system. As an example, one might have thought in the case of converting one particle into two that the probability rate of reverse back to one particle ought to be 2δ, with a resulting bound on entropy production of the form $\ln(g/\delta) - \ln 2$. To see why the $\ln 2$ ought not to be included, it is important to recognize that particles in a classical physical system have locations that distinguish them from each other. Thus, we either can think of the process in question as bounding the entropy production of self-replication *plus* mixing of particles (in which case the mixing entropy term cancels the factor of 2) or else we can define our coarse-grained states in the transition in question so that we only consider the probability current for the reversion of the replicator that was just born. Either way, the relevant bound comes out to $\ln(g/\delta)$.

The above result is certainly consistent with our expectation that in order for this self-replicator to be capable of net growth, we have to require that $g > \delta$, which in turn sets a *positive* lower bound on the total entropy production associated with self-replication. More can be seen, however, by rearranging the expression at fixed δ, Δs_{int}, and Δq, to obtain

$$g_{max} - \delta = \delta\left(\exp[\beta\Delta q + \Delta s_{int}] - 1\right). \quad (11)$$

In other words, the maximum net growth rate of a self-replicator is fixed by three things: its internal entropy (Δs_{int}), its durability ($1/\delta$), and the heat that is dissipated into the surrounding bath during the process of replication Δq.

Several comments are in order. First of all, let us consider comparing two different replicators with fixed Δs_{int} and δ that have different heat values Δq and $\Delta q'$. If $\Delta q > \Delta q'$, then clearly $g_{max} > g'_{max}$; the replicator that dissipates more heat has the potential to grow accordingly faster. Moreover, we know by conservation of energy that this heat has to be generated in one of two different ways: either from energy initially stored in reactants out of which the replicator gets built (such as through the hydrolysis of sugar) or else from work done on the system by some time-varying external driving field (such as through the absorption of light during photosynthesis). In other words, basic thermodynamic constraints derived from exact considerations in statistical physics tell us that a self-replicator's maximum potential fitness is set by how effectively it exploits sources of energy in its environment to catalyze its own reproduction. Thus, the empirical, biological fact that reproductive fitness is intimately linked to efficient metabolism now has a clear and simple basis in physics.

A somewhat subtler point to make here is that self-replicators can increase their maximum potential growth rates not only by more effectively fueling their growth via Δq, but also by lowering the cost of their growth via δ and Δs_{int}. The less durable or the less organized a self-replicator is, all things being equal, the less metabolic energy it must harvest at minimum in order to achieve a certain growth rate. Thus, in a competition among self-replicators to dominate the population of the future, one strategy for "success" is to be simpler in

construction and more prone to spontaneous degradation. Of course, in the limit where two self-replicators differ dramatically in their internal entropy or durability (as with a virus and a rabbit) the basis for comparison in terms of Darwinian fitness becomes too weak. Nevertheless, in a race between competitors of similar form and construction, it is worthwhile to note that one strategy for reducing the minimal "metabolic" costs of growth for a replicator is to substitute in new components that are likely to "wear out" sooner.

SELF-REPLICATING POLYNUCLEOTIDES

A simple demonstration of the role of durability in replicative fitness comes from the case of polynucleotides, which provide a rich domain of application for studying the role of entropy production in the dynamics of self-replication.[8] One recent study has used *in vitro* evolution to optimize the growth rate of a self-replicating RNA molecule whose formation is accompanied by a single backbone ligation reaction and the leaving of a single pyrophosphate group.[9] It is reasonable to assume that the reverse of this reaction, in which the highly negatively charged pyrophosphate would have to attack the phosphate backbone of the nucleic acid, would proceed more slowly than a simple hydrolysis by water. The rate of such hydrolysis has been measured by carrying out a linear fit to the early-time (20 day) exponential kinetics of spontaneous degradation for the RNAase A substrate UpA, and the half-life measured is on the order of 4 years.[10] Thus, with a doubling time of 1 h for the self-replicator, we can estimate for this system that $\ln[g/\delta] \geq \ln[(4 \text{ years})/(1 \text{ h})]$. Since the ligation reaction trades the mixing entropy of substrate for that of pyrophosphate at comparable ambient concentrations, we can also assume in this case that the change in internal entropy for the reaction is negligible. Thus, we can estimate the heat bound as

$$\langle \Delta Q \rangle \geq RT \ln[(4 \text{ years})/(1 \text{ h})] = 7 \text{ kcal mol}^{-1}. \quad (12)$$

Since experimental data indicate an enthalpy for the reaction in the vicinity of 10 kcal mol^{-1},[11,12] it would seem this molecule operates quite near the limit of thermodynamic efficiency set by the way it is assembled.

To underline this point, we may consider what the bound might be if this same reaction were somehow achieved using DNA, which in aqueous solution is much more kinetically stable against hydrolysis than RNA.[13] In this case, we would have $\langle \Delta Q \rangle \geq RT \ln[(3 \times 10^7 \text{ years})/(1 \text{ h})] = 16 \text{ kcal mol}^{-1}$, which exceeds the estimated enthalpy for the ligation reaction and is therefore prohibited thermodynamically. This calculation illustrates a significant difference between DNA and RNA, regarding each molecule's ability to participate in self-catalyzed replication reactions fueled by simple triphosphate building blocks: the far greater durability of DNA demands that a much higher per-base thermodynamic cost be paid in entropy production[14] in order for for the growth rate to match that of RNA in an all-things-equal comparison.

The key point here is that if a self-replicating nucleic acid catalyzes its own growth with a rate constant g and is degraded with rate constant δ, then the molecule should be capable of exhibiting exponential growth through self-replication with a doubling time proportional to $1/(g - \delta)$. However, thermodynamics only sets a bound on the ratio g/δ, which means that $g - \delta$ can be made arbitrarily large while the "metabolic cost" of replication remains fixed. Thus, surprisingly, the greater fragility of RNA could be seen as a fitness advantage in the right circumstances, and this applies even if the system in question is externally driven, and even still if the replicators maintain non-Boltzmann distributions over their internal degrees of freedom. Moreover, we would expect that the heat bound difference between DNA and RNA should increase roughly linearly in the number of bases ℓ ligated during the reaction, which forces the maximum possible growth rate for a DNA replicator to shrink *exponentially* with ℓ in comparison to that of its RNA competitor in an all things equal comparison. This observation is certainly intriguing in light of past arguments made on other grounds that RNA, and not DNA, must have acted as the material for the pre-biotic emergence of self-replicating nucleic acids.[12,15]

The particular example of a single nucleic acid base ligation considered here is instructive because it is a case where the relationship between irreversibility and entropy production further simplifies into a recognizable form. It is reasonable to define our coarse-graining of states so that the starting and ending products of this reaction are each in a local, thermal equilibrium, such that all of their degrees of freedom are Boltzmann-distributed given the constraints of their chemical bonds. In such a circumstance, detailed balance alone requires that the ratio of the forward and reverse rates be locked to the Gibbs' free energy change for the reaction via the familiar formula $\ln[k_f/k_r] = -\beta \Delta G \equiv -\beta \Delta H + \Delta S_{int} = \beta \Delta Q + \Delta S_{int}$.[8] In this light, the relationship between durability and growth rate has an elegant description in terms of transition state theory: an activation barrier that is lower in the forward direction will be lower in the reverse direction as well. The key point here, however, is that whereas the relationship between free energy and reaction rates obtains only under local equilibrium assumptions, the inequality we have derived here bounding entropy production in terms of irreversibility applies even in cases where many degrees of freedom in the system start and end the replication process out of thermal equilibrium.

On a related note, it may also be pointed out that a naïve objection might be raised to the above account of nucleic acid growth on the grounds that its conclusions about the maximum possible growth rate of DNA would seem to disagree with the empirical fact that DNA is obviously capable of undergoing replication much more rapidly than one phosphodiester linkage per hour; the processive holoenzyme DNA polymerase III is well-known for its blazing catalytic speed of ~1000 base-pairs per second.[16] The resolution of this puzzle lies in realizing that in the protein-assisted DNA replication scenario being raised here, the polymerase assembly is first loaded onto DNA in an ATP-dependent manner that irreversibly attaches the enzyme to the strand using a doughnut-shaped clamp composed of various protein subunits. Thus, while the polymerase can catalyze the elongation of the DNA chain extremely rapidly, one also must take into account that the *reverse* reaction of DNA hydrolysis should happen much

more readily with an enzyme tethered to the strand than it does for isolated DNA in solution. This example therefore underlines the care that must be shown in how the coarse-grained states **I** and **II** are defined in the computation of irreversibility.

BACTERIAL CELL DIVISION

Ligation of polynucleotides provides a relatively simple test case where tracking the formation of a new replicator can be reduced to monitoring the progress of a single molecular event (i.e., the formation of a phosphodiester linkage); however, the macroscopic relationship between irreversibility and entropy production applies equally in cases that are much more complex. Indeed, we shall now argue that these basic thermodynamic constraints are even relevant to the growth and division of whole single-celled organisms.

We begin by considering the preparation of a large system initially containing a single *E. coli* bacterium, immersed in a sample of rich nutrient media in contact with a heat bath held at the bacterial cell's optimal growth temperature ($1/\beta \equiv T \sim 4.3 \times 10^{-21}$ J).[17, 18] We can assume furthermore that the cell is in exponential growth phase at the beginning of its division cycle, and that, while the volume and mass of the entire system are held fixed so that no particles are exchanged with any external bath and the walls of the container do not move, the composition and pressure of the nutrient media surrounding the bacterium mimics that of a well-aerated sample open to the earth's atmosphere. If we summarize the experimental conditions described above with the label **I**, we can immediately say that there is some probability $p(i|\mathbf{I})$ that the system is found in some particular microstate i (with energy E_i) given that it was prepared in the macroscopic condition **I** by some standard procedure. Although this probability might well be impossible to derive *ab initio*, in principle it could be measured (at least in a thought-experiment) through lengthy experimental consultation with a microbiologist.

Three preliminary points are in order as a result of the fact that the "system" being defined here is not only the bacterium, but also the surrounding broth containing the food it will eat and the oxygen it will breath, etc. First of all, looking ahead to our expression for entropy production, we must consider the entropy changes associated with the bacterium's ingestions and excretions during growth and division to be part of the *internal* entropy change of the system ΔS_{int}. The somewhat unexpected result is that if, for example, the bacterium subsisted on metabolic reactions that were accompanied by large increases in entropy (through, for example, anaerobic gas evolution[19]), the total internal entropy change for forming a new bacterium could actually be positive.

Second, it must be pointed out that the connection between the internal Shannon entropy $S_{int} \equiv -\overline{\ln p} \equiv -\sum p \ln p$ and the familiar idea we have of entropy from equilibrium statistical mechanics is not obvious, and requires elaboration. In Boltzmann-distributed systems, it has long been clear that the internal Shannon entropy has a natural connection to the heat exchanged with the surroundings because of the set relationship between internal energy and probability. Only more recently has it been demonstrated that S_{int}

continues to obey a general thermodynamic relation for arbitrary nonequilibrium transitions: $\Delta S_{int} \leq \Delta Q_{ex}/T$. Here, ΔQ_{ex} is the so-called excess heat, which measures the extra heat evolution on top of what would accompany the various dissipative steady-states traversed during the transition.[20] Perhaps more importantly, we expect that even far from equilibrium, the Shannon entropy remains the measure of "statistical disorder" in the system that it is for equilibrium systems. The easiest way to see this is consider a case of uniform starting and ending distributions, that is, where $p(i|\mathbf{I}) = p_{\mathbf{I}}$ and $p(j|\mathbf{II}) = p_{\mathbf{II}}$ for all states i and j that have non-zero probability in their respective ensembles. Clearly, in this case $\Delta S_{int} = \ln [p_{\mathbf{I}}/p_{\mathbf{II}}] = \ln [\Omega_{\mathbf{II}}/\Omega_{\mathbf{I}}]$, that is, it simply measures how many more or fewer states there are in **II** than in **I**. Moreover, it is more generally the case that the Shannon entropy is effecting a log-scale comparison of volumes in phase space before and after the transition. Accordingly, the factors affecting this quantity far from equilibrium should be the same as they are near equilibrium, namely the number of different possible positions and velocities available to particles in the system when they are arranged to belong to the ensemble in question. Thus changes in partial volumes of gases are, for example, still relevant to the question of how this internal entropy has changed.

Finally, it should be noted that for the case of a microbe such as *E. coli* (which is not carrying out any sort of externally driven photosynthetic process) the propagator for the system $\pi(i \rightarrow j; \tau)$ over any interval of time τ can be taken to obey (2). Put another way, while the bacterium by itself might appear to be driven by external currents of chemical reactants and products, the system as a whole (which includes the nutrient media) is not driven at all, and simply exchanges heat with a surrounding thermal reservoir. Thus, quite counterintuitively, we expect that the eventual steady-state for this system (as for any system of fixed volume and particle number left in contact with a heat bath for an infinite amount of time) will be a Boltzmann distribution over microstates in which detailed balance holds. What makes the scenario of interest here a far-from-equilibrium process, then, is only that its initial conditions $p(i|\mathbf{I})$ correspond to a *highly* non-Boltzmann distribution over microstates. The bacterial growth that takes place at the very beginning of the long process of relaxation to equilibrium is in this sense a mere transient, the fuel for which comes from chemical energy stored in the system's starting microscopic arrangement. By the time detailed balance sets in, we expect all bacteria initially in the system will have long since overgrown their container, starved, perished, and disintegrated into their constituent parts.

Now consider what would happen in our system if we started off in some microstate i in **I** and then allowed things to propagate for a time interval of τ_{div}, the typical duration of a single round of growth and cell division. From the biological standpoint, the expected final state for the system is clear: two bacteria floating in the media instead of one, and various surrounding atoms rearranged into new molecular combinations (e.g., some oxygen converted into carbon dioxide). We can label the ensemble of future states corresponding to such a macroscopic outcome **II**. Given, that the system was initially prepared in **I** and did subsequently end up in **II**, any

microstate j will have some finite likelihood which we can call $p(j|\mathbf{II})$.

For the process of bacterial cell division introduced above, our ensemble $\mathbf{II}$ is a bath of nutrient-rich media containing two bacterial cells in exponential growth phase at the start of their division cycles. In order to make use of the relation in (8), we need to estimate $\pi(\mathbf{II} \rightarrow \mathbf{I})$, the likelihood that after time τ_{div}, we will have ended up in an arrangement $\mathbf{I}$ where only one, newly formed bacterium is present in the system and another cell has somehow been converted back into the food from which it was built. Of course, cells are never observed to run their myriad biochemical reactions backwards, any more so than ice cubes are seen forming spontaneously in pots of boiling water. Nevertheless, it is implicit in the assumptions of very well-established statistical mechanical theories that such events have non-zero (albeit absurdly small) probabilities of happening,[1] and these likelihoods can be bounded from above using the probability rates of events we *can* measure experimentally. This is possible because the rough physical features of the system are sufficient for making plausible estimates of thermodynamic quantities of interest; since we are ultimately interested in bounding the heat generated by this process, we are only concerned with the impact of the probabilities we estimate on a logarithmic scale; we would have to change our probability estimate by many orders of magnitude to see any effect on the corresponding heat bound.

The first piece is relatively easy to imagine: while we may not be able to compute the exact probability of a bacterium fluctuating to peptide-sized pieces and de-respiring a certain amount of carbon dioxide and water, we can be confident it is *less* likely than all the peptide bonds in the bacterium spontaneously hydrolyzing. Happily, this latter probability may be estimated in terms of the number of such bonds n_{pep}, the division time τ_{div}, and the peptide bond half-life τ_{hyd}.

An added complication comes from the fact that the bacterium normally grows at a rate of $r = n_{pep}/\tau_{div}$ peptide bonds formed per unit time.[18] In order to make a model for the probability of a bacterium disintegrating over a set period of time, we need to specify when during that period it starts to disintegrate, and here two different factors collide. The less time we wait for the disintegration to begin, the less probable it is that it will happen in that period of time. However, the longer we wait, the more new peptide bonds will get synthesized in the normal course of growth, so that the probability of subsequently undergoing spontaneous disintegration of the whole cell in the remainder of the time τ_{div} shrinks exponentially. Between these two countervailing effects there must be an optimal time that maximizes the probability of disintegration; thus, assuming each peptide bond spontaneously hydrolyzes independently from the others with probability rate $\sim \tau_{hyd}^{-1}$, we may therefore model the total cell hydrolysis probability p_{hyd} in time t as

$$\ln p_{hyd} \simeq (n_{pep} + rt)\ln[t/\tau_{hyd}]. \quad (13)$$

This quantity is maximized for t_{max} satisfying $\tau_{div}/t_{max} + \ln[t_{max}/\tau_{hyd}] + 1 = 0$, which we can compute numerically for chosen values of the two input timescales. Following this,

we simply have to evaluate

$$|\ln p_{hyd}| \simeq |n_{pep} \ln[t_{max}/\tau_{hyd}]| = n_{pep}(\tau_{div}/t_{max} + 1). \quad (14)$$

We have now dealt with the demise of one of the cells. Handling the one that stays alive is more challenging, as we have assumed this cell is growing processively, and we ought not make the mistake of thinking that such a reaction can be halted or paused by a small perturbation. The onset of exponential growth phase is preceded in *E. coli* by a lag phase that can last several hours,[17] during which gene expression is substantially altered so as to retool the cell for rapid division fueled by the available metabolic substrates.[21] It is therefore appropriate to think of the cell in question as an optimized mixture of components primed to participate in irreversible reactions like nutrient metabolism and protein synthesis.

We can therefore argue that the likelihood of a spontaneous, sustained pause (of duration τ_{div}) in the progression of these reactions is very small indeed: if each enzymatic protein component of the cell were to reject each attempt of a substrate to diffuse to its active site (assuming a diffusion time of small molecules between proteins of $\tau_{diff} \sim 10^{-8}$ s[22,23]), we would expect $|\ln p_{pause}| \propto |n_{pep}(\tau_{div}/\tau_{diff})|$ to exceed $|\ln p_{hyd}|$ by orders magnitude. We must, however, consider an alternative mechanism for the most likely $\mathbf{II} \rightarrow \mathbf{I}$ transition: it is possible that a cell could grow and divide in an amount of time slightly less than τ_{div}.[18] If, subsequent to such an event, the daughter cell of the recent division were to spontaneously disintegrate back into its constituent nutrients (with log-probability at most on the order of $\ln p_{hyd}$), we would complete the interval of τ_{div} with one, recently divided, processively growing bacterium in our system, that is, we would have returned to the $\mathbf{I}$ ensemble. Thus, via a back-door into $\mathbf{I}$ provided to us by bacterial biology, we can claim that

$$\ln \pi(\mathbf{II} \rightarrow \mathbf{I}) \leq 2\ln p_{hyd} \simeq -2n_{pep}(\tau_{div}/t_{max} + 1). \quad (15)$$

Having obtained the above result, we can now refer back to the bound we set for the heat produced by this self-replication process and write

$$\beta\langle Q\rangle \geq 2n_{pep}(\tau_{div}/t_{max} + 1) - \Delta S_{int}. \quad (16)$$

This relation demonstrates that the heat evolved in the course of the cell making a copy of itself is set not only by the decrease in entropy required to arrange molecular components of the surrounding medium into a new organism, but also by how rapidly this takes place (through the division time τ_{div}) and by how long we have to wait for the newly assembled structure to start falling apart (through t_{max}). Moreover, we can now quantify the extent of each factor's contribution to the final outcome, in terms of n_{pep}, which we estimate to be 1.6×10^9, assuming the dry mass of the bacterium is 0.3 picograms.[24]

The total amount of heat produced in a single division cycle for an *E. coli* bacterium growing at its maximum rate on lysogeny broth (a mixture of peptides and glucose) is $\beta\langle Q\rangle = 220n_{pep}$.[17] We expect the largest contributions to the internal entropy change for cell division to come from the equimolar conversion of oxygen to carbon dioxide (since carbon dioxide has a significantly lower partial pressure in the

atmosphere), and from the confinement of amino acids floating freely in the broth to specific locations inside bacterial proteins. We can estimate the contribution of the first factor (which increases entropy) by noting that $\ln(\upsilon_{CO_2}/\upsilon_{O_2}) \sim 6$. The liberation of carbon from various metabolites also increases entropy by shuffling around vibrational and rotational degrees of freedom, but we only expect this to make some order unity modification to the entropy per carbon atom metabolized. At the same time, peptide anabolism reduces entropy: by assuming that in 1% tryptone broth, an amino acid starts with a volume to explore of $\upsilon_i = 100$ nm^3 and ends up tightly folded up in some $\upsilon_f = 0.001$ nm^3 sub-volume of a protein, we obtain $\ln(\upsilon_f/\upsilon_i) \sim -12$. In light of the fact that the bacterium consumes during division a number of oxygen molecules roughly equal to the number of amino acids in the new cell it creates,[17,25] we can arbitrarily set a generous upper bound of $-\Delta S_{int} \leq 10 n_{pep}$.

In order to compare this contribution to the irreversibility term in (8), we assume a cell division time of 20 min,[17,18] and a spontaneous hydrolysis lifetime for peptide bonds of $\tau_{hyd} \sim 600$ years at physiological pH,[26] which yields a $t_{max} \sim 1$ min and $2 n_{pep}(\tau_{div}/t_{max} + 1) = 6.7 \times 10^{10} \simeq 42 n_{pep}$, a quantity at least several times larger than ΔS_{int}. We often think of the main entropic hurdle that must be overcome by biological self-organization as being the cost of assembling the components of the living thing in the appropriate way. Here, however, we have evidence that this cost for aerobic bacterial respiration is relatively small,[19] and is substantially outstripped by the sheer irreversibility of the self-replication reaction as it churns out copies that do not easily disintegrate into their constituent parts.

More significantly, these calculations also establish that the *E. coli* bacterium produces an amount of heat less than six times $(220 n_{pep}/42 n_{pep})$ as large as the absolute physical lower bound dictated by its growth rate, internal entropy production, and durability. In light of the fact that the bacterium is a complex sensor of its environment that can very effectively adapt itself to growth in a broad range of different environments, we should not be surprised that it is not perfectly optimized for any given one of them. Rather, it is remarkable that in a single environment, the organism can convert chemical energy into a new copy of itself so efficiently that if it were to produce even a quarter as much heat it would be pushing the limits of what is thermodynamically possible! This is especially the case since we deliberately underestimated the reverse reaction rate with our calculation of p_{hyd}, which does not account for the unlikelihood of spontaneously converting carbon dioxide back into oxygen. Thus, a more accurate estimate of the lower bound on $\beta\langle Q \rangle$ in future may reveal *E. coli* to be an even more exceptionally well-adapted self-replicator than it currently seems.

At the same time, it is worthwhile to remember that any self-replicator faces other constraints in addition to the thermodynamic ones explored in the preceding analysis. Any entity that makes copies of itself is built out of some particular set of materials with a contingent set of properties and modes of interaction or intercombination. Viewed purely from the standpoint of the generalization of the Second Law of Thermodynamics derived in this work, it is possible that *E. coli*

might grow dramatically faster than it does currently while exhausting the same amount of heat. Once one accounts for the contingencies of exactly how bacteria replicate (and, for example, how they use ribosomes to accurately synthesize the protein machinery that allow them to function at the molecular level[27]), it is reasonable to expect that the organism in fact faces far more restrictive impediments to growth that arise from the speeds of rate-limiting reactions on which the whole process depends.

CLOSING REMARKS

The process of cellular division, even in a creature as ancient and streamlined as a bacterium, is so bewilderingly complex that it may come as some surprise that physics can make any binding pronouncements about how fast it all can happen. The reason this becomes possible is that nonequilibrium processes in constant temperature baths obey general laws that relate forward and reverse transition probabilities to heat production.[2] Previously, such laws had been applied successfully in understanding thermodynamics of copying "informational" molecules such as nucleic acids.[8] In those cases, however, the information content of the system's molecular structure could more easily be taken for granted, in light of the clear role played by DNA in the production of RNA and protein.

What we have glimpsed here is that the underlying connection between entropy production and transition probability has a much more general applicability, so long as we recognize that "self-replication" is only visible once an observer decides how to classify the "self" in the system: only once a coarse-graining scheme determines how many copies of some object are present for each microstate can we talk in probabilistic terms about the general tendency for that type of object to affect its own reproduction, and the same system's microstates can be coarse-grained using any number of different schemes. Whatever the scheme, however, the resulting stochastic population dynamics must obey the same general relationship entwining heat, organization, and durability. We may hope that this insight spurs future work that will clarify the general physical constraints obeyed by natural selection in nonequilibrium systems.

[1]C. W. Gardiner, *Handbook of Stochastic Methods*, 3rd ed. (Springer, 2003).
[2]G. E. Crooks, Phys. Rev. E **60**, 2721 (1999).
[3]R. A. Blythe, Phys. Rev. Lett. **100**, 010601 (2008).
[4]A. Gomez-Marin, J. M. Parrondo, and C. Van den Broeck, Phys. Rev. E **78**, 011107 (2008).
[5]G. Verley, R. Chétrite, and D. Lacoste, Phys. Rev. Lett. **108**, 120601 (2012).
[6]R. Landauer, IBM J. Res. Dev. **5**, 183 (1961).
[7]M. A. Nowak, Trends Ecol. Evol. **7**, 118 (1992).
[8]D. Andrieux and P. Gaspard, Proc. Natl. Acad. Sci. U.S.A. **105**, 9516 (2008).
[9]T. A. Lincoln and G. F. Joyce, Science **323**, 1229 (2009).

[10] J. E. Thompson, T. G. Kutateladze, M. C. Schuster, F. D. Venegas, J. M. Messmore, and R. T. Raines, Bioorg. Chem. **23**, 471 (1995).

[11] C. A. Minetti, D. P. Remeta, H. Miller, C. A. Gelfand, G. E. Plum, A. P. Grollman, and K. J. Breslauer, Proc. Natl. Acad. Sci. U.S.A. **100**, 14719 (2003).

[12] H. J. Woo, R. Vijaya Satya, and J. Reifman, PLoS Comput. Biol. **8**, e1002534 (2012).

[13] G. K. Schroeder, C. Lad, P. Wyman, N. H. Williams, and R. Wolfenden, Proc. Natl. Acad. Sci. U.S.A. **103**, 4052 (2006).

[14] D. Y. Zhang, A. J. Turberfield, B. Yurke, and E. Winfree, Science **318**, 1121 (2007).

[15] W. Gilbert, Nature (London) **319**, 618 (1986).

[16] Z. Kelman and M. O'Donnell, Annu. Rev. Biochem. **64**, 171 (1995).

[17] H. P. Rothbaum and H. M. Stone, J. Bacteriol. **81**, 172 (1961).

[18] P. Wang, L. Robert, J. Pelletier, W. L. Dang, F. Taddei, A. Wright, and S. Jun, Curr. Biol. **20**, 1099 (2010).

[19] U. von Stockar, T. Maskow, J. Liu, I. W. Marison, and R. Patino, J. Biotechnol. **121**, 517 (2006).

[20] T. Hatano and S. Sasa, Phys. Rev. Lett. **86**, 3463 (2001).

[21] D. E. Chang, D. J. Smalley, and T. Conway, Mol. Microbiol. **45**, 289 (2002).

[22] D. Brune and S. Kim, Proc. Natl. Acad. Sci. U.S.A. **90**, 3835 (1993).

[23] S. C. Blacklow, R. T. Raines, W. A. Lim, P. D. Zamore, and J. R. Knowles, Biochemistry **27**, 1158 (1988).

[24] F. C. Neidhardt, *E. coli and Salmonella: Cellular and Molecular Biology* (ASM Press, 1990), Vol. 1.

[25] C. L. Cooney, D. I. Wang, and R. I. Mateles, Biotechnol. Bioeng. **11**, 269 (1969).

[26] A. Radzicka and R. Wolfenden, J. Am. Chem. Soc. **118**, 6105 (1996).

[27] C. G. Kurland and M. Ehrenberg, Annu. Rev. Biophys. Biophys. Chem. **16**, 291 (1987).