Theory of Molecular Machines.
II. Energy Dissipation from Molecular Machines

running title: Dissipation from Molecular Machines

Thomas D. Schneider*

version = 2.13 of edmm.tex 2004 Feb 3
Version 2.01 was submitted 1990 December 5
Schneider, T. D. (1991*b*). Theory of molecular machines. II. Energy dissipation from molecular machines. *J. Theor. Biol.* **148**, 125-137.
http://www.lecb.ncifcrf.gov/˜toms/paper/edmm

Single molecules perform a variety of tasks in cells, from replicating, controlling and translating the genetic material to sensing the outside environment. These operations all require that specific actions take place. In a sense, each molecule must make tiny decisions. To make a decision, each "molecular machine" must dissipate an energy P_y in the presense of thermal noise N_y. The number of binary decisions that can be made by a machine which has d_{space} independently moving parts is the "machine capacity" $C_y = d_{space} \log_2(\frac{P_y+N_y}{N_y})$. This formula is closely related to Shannon's channel capacity for communications systems, $C = W \log_2(\frac{P+N}{N})$.

This paper shows that the minimum amount of energy that a molecular machine must dissipate in order to gain one bit of information is $\mathcal{E}_{min} =$

*National Cancer Institute, Frederick Cancer Research and Development Center, Laboratory of Experimental and Computational Biology, National Cancer Institute, P. O. Box B, Building 144, Room 469, Frederick, MD 21702. email address: toms@ncifcrf.gov.

$k_BT\ln(2)$ joules per bit. This equation is derived in two distinct ways. The first derivation begins with the Second Law of Thermodynamics, which shows that the statement that there is a minimum energy dissipation is a restatement of the Second Law of Thermodynamics. The second derivation begins with the machine capacity formula, which shows that the machine capacity is also related to the Second Law of Thermodynamics.

One of Shannon's theorems for communications channels is that as long as the channel capacity is not exceeded, the error rate may be made as small as desired by a sufficiently involved coding. This result also applies to the dissipation formula for molecular machines. So there is a *precise* upper bound on the number of choices a molecular machine can make for a given amount of energy loss. This result will be important for the design and construction of molecular computers.

Introduction

The relationship between entropy and living things has been widely discussed since the last century. In 1871, Maxwell unnerved thermodynamicists by suggesting a way that a living being could break the Second Law of Thermodynamics[1,2,3]. One of the many implications of the Second Law[4] is that when a quantity of gas is separated into two compartments, both initially at the same pressure and temperature, it is not possible to raise the temperature in one of the compartments and lower the other without performing work. Maxwell proposed that a tiny intelligent being could open and close a hole between two such compartments to allow only fast molecules from the first compartment to pass into the second compartment and slow molecules to move in the other direction. The first compartment would become cool while the other compartment would become hot. Assuming that this demon does not need to do any work to open and close the hole, one could use the heat difference to run an engine. This perpetual motion machine would violate the Second Law of Thermodynamics.

The problem of Maxwell's demon was partially resolved by Szilard in 1929 [5,6,7] and more completely by Brillouin in 1951[8,9]. They recognized that the demon would have to obtain one bit of information about the approaching molecules. To distinguish the molecules from the background of thermal radiation, the demon could use a flashlight. Brillouin showed that more energy would be lost by operating the flashlight than could be gained by the demon's tricks. Thus the information that the demon gains must be paid for by a loss of some energy, and the Second Law is not broken.

Brillouin and Szilard's arguments are not convincing because the problem has been posed for an imaginary beast. It is not obvious, for example, that *controlled* opening and closing of a door can be done without energy dissipation. The difficulty of guaranteeing that a photon from the flashlight reaches the eye of the demon and the problem of what happens to the photon's energy in the eye of the demon have also been ignored. To bring this problem into the concrete world of molecular biology[10], we can focus on the mechanisms of molecules that can be investigated in the laboratory. The question of what Maxwell's demon can do becomes a question of how rhodopsin in the eye and actomyosin in muscle operate. Indeed, it becomes the question of how all molecular machines operate.

Molecular Machines A molecular machine is a single macromolecule or macromolecular complex that performs a specific function for a living system [11]. For example, single stranded DNA can hybridize to form duplex DNA[12]. This *operation* is defined by two limiting states, *before* the operation when the strands are separated, and *after* the operation when complementary strands are paired. Consider the First Law of Thermodynamics for this operation:

$$\Delta U = q + w \tag{1}$$

where ΔU is the change in internal energy, q is the heat flowing into the machine and w is the work done by the machine on the surroundings. (The defined directions of q and w reflect their original use to describe the input of heat and extraction of work from steam engines.) Since the DNA molecule does not do work on an external object when it hybridizes, $w = 0$. The internal energy of the machine decreases, $\Delta U < 0$, so heat is dissipated into the surroundings, $q < 0$. How can we characterize the action that the machine has taken if it does not do external work? Although the operation can be characterized by the energy dissipated, the important biological aspect of the operation is the number of choices that the machine makes. Thus, to form each base pair of DNA, only 4 out of 16 possibilities are acceptable. This 1 in 4 choice represents $\log_2 4 = 2$ bits of information "gained" by the machine. Other examples and a detailed definition of molecular machines and their operations are given in[11].

Overview of the Derivations In this paper I derive a formula that relates energy to information in the context of molecular machines. Before the formula can be derived, it is necessary to define information and to distinguish this definition from others that appear in the literature.

The formula for information shows that bits on the microscopic level are conceptually the same as bits on the macroscopic level. The formula allows us to determine the quantitative relationship between information and entropy, a topic which has led to much confusion in the literature.

The minimum energy that must be dissipated in order to gain one bit of information, $\mathcal{E}_{min} = k_B T \ln(2)$ joules per bit, is first derived from the Second Law of Thermodynamics. The derivation is straight forward, given the definition of information, but to my knowledge it does not appear in the literature.

The same formula for $\mathcal{E}_{min}$ is then derived from the machine capacity formula, equation (17).

I then show that molecular machines perform precise logical operations. This implies that computers made from single molecules are possible. Such computers should be able to approach the ideal minimum energy dissipation.

Uncertainty, Entropy, and Information

Suppose that a molecular machine has Ω possible microstates, each with a particular probability P_i:

$$\sum_{i=1}^{\Omega} P_i = 1 \quad \text{and} \quad P_i \geq 0. \tag{2}$$

The set of all possible microstates forms a sphere in a high dimensional space[11], and Ω is proportional to the volume of the sphere[13]. Each "microstate" represents a particular machine configuration. We may write the uncertainty of the machine's microstates using Shannon's formula[14,15,16,17]:

$$H \equiv -\sum_{i=1}^{\Omega} P_i \log_2 P_i \quad \text{(bits per microstate)}. \tag{3}$$

Likewise, the Boltzmann-Gibbs entropy of a physical system, such as a molecular machine, is

$$S \equiv -k_B \sum_{i=1}^{\Omega} P_i \ln P_i \quad \left(\frac{\text{joules}}{\text{K} \cdot \text{microstate}}\right) \tag{4}$$

where k_B is Boltzmann's constant $(1.38 \times 10^{-23}$ joules / K$)$[18,19]. Since $\log_2(x) = \ln(x)/\ln(2)$,

$$S = k_B \ln(2) \left[-\sum_{i=1}^{\Omega} P_i \log_2 P_i\right]. \tag{5}$$

Substituting equation (3) into (5) gives

$$S = k_{\mathrm{B}} \ln(2) H. \tag{6}$$

The only difference between uncertainty and entropy for the microstates of a macromolecule is in the units of measure, bits versus joules per K respectively [20,21,3].

The entropy of a molecular machine may decrease at the expense of a larger increase of entropy in the surroundings. For a decrease in the entropy during a machine operation:

$$\Delta S = S_{after} - S_{before} \quad \left(\frac{\text{joules}}{\text{K} \cdot \text{operation}}\right) \tag{7}$$

there is a corresponding decrease in the uncertainty of the machine:

$$\Delta H = H_{after} - H_{before} \quad \text{(bits per operation)}. \tag{8}$$

Using (6) we find:

$$\Delta S = k_{\mathrm{B}} \ln(2) \Delta H. \tag{9}$$

When the uncertainty of a machine decreases during an operation, it gains some information R [14,9,22,3] defined by:

$$R \equiv -\Delta H \quad \text{(bits per operation)}. \tag{10}$$

This is the information discussed in [11] and measured in [23]. It is important to notice that H_{after} is not always zero. For example, a DNA sequence recognizer may accept a purine at some position in a binding site, in which case H_{after} is 1 bit. Thus we cannot equate information gained (R) with the uncertainty before an operation takes place (H_{before}) nor with the uncertainty remaining after an operation has been completed (H_{after}). Use of definition (10) avoids a good deal of confusion found in the literature [24,25,26,27,28].

In particular, the largest possible value of R is obtained when H_{after} is as small as possible (perhaps close to zero) and H_{before} is maximized. The latter occurs only when the symbols are equally likely, in which case equation (3) collapses to $H_{equal} = \log_2 \Omega$. In the same way, if there are M_y symbols, the information required to chose one of them is $\log_2 M_y$. This form was used by Shannon [29] and in the previous paper of this series to determine the capacity formulas.

Substituting (10) into (9) gives:

$$\Delta S = -k_{\mathrm{B}} \ln(2) R. \tag{11}$$

This equation gives a direct, quantitative relationship between the decrease in entropy of a molecular machine and the information that it gains during an operation [3]. We must carefully note that ΔS in (11) refers only to that *part* of the total entropy change that accounts for the selection of states made by the machine during an operation. Since R is positive for an operation, this ΔS is always negative. For the operation to proceed, the total entropy of the universe must increase or remain the same:

$$\Delta S_{universe} = \Delta S + \Delta S_{surround} \geq 0. \quad (12)$$

For example, the equality holds when a solution of *Eco*RI and DNA is at equilibrium in the absence of Mg^{++} to prevent cutting. Priming and machine operations occur, but the entropy of the universe does not increase. In other words, the entropy of the *local* surroundings must increase in compensation for a molecular machine's entropy decrease during an operation, $\Delta S_{surround} \geq -\Delta S$.

Other Definitions of Information Do Not Apply to Molecular Machines

The formulation for R accounts for a single molecular machine either gaining or losing information as it cycles through its operations [11]. A similar formula, $I = \sum_{i=1}^{\Omega} P_{i,after} log_2(P_{i,after}/P_{i,before})$ [30,31] gives the maximum information an *observer* could gain by observing a system. I is always zero or positive [30]. If we were to start a molecular machine in some state A, and we later observe it in another state B ($P_{i,A} \neq P_{i,B}$, for some i) then $I_{AB} > 0$. If the machine returns to A, then $I_{BA} > 0$, so $I_{AB} + I_{BA} > 0$, meaning that the observer learned the details of how the machine performed this cycle. *But the machine itself is in the same state as it began, so it cannot have gained any information*, just as a computer memory does not gain any information if we fill it with data and then remove the data again. Thus *only a path independent function of state, such as R, is appropriate to use for the information a single molecular machine gains during its operation.* External observers and the measurements they may make are not relevant to the problem.

How Uncertainty Decreases Define Information

By using a decrease in uncertainty to define information (equation (10)), we also avoid dealing with absolute quantities. Information is gained when a machine changes from an indeterminate state to a more determined state [11]. There are a large number of microstates in both the *before* and *after* states, but since we are only concerned with changes of state, the large numbers are removed from consideration when the subtraction is made. Although this might appear to be a difference between large numbers, it is not: it is the *logarithm of the ratio* of large numbers (the sphere volumes in[11]), which can be quite small. Because of this we can even legitimately speak about single bits for changes in a macroscopic object without knowing the detailed state of its molecules. Consider a coin flipping in the air. The entropy of this system is enormous, on the order of $H_{before} = 10^{23}$ bits in a 3 gram copper penny at 300K using equation (6) and data from[19]. If all states were equally likely, $P_i = \frac{1}{\Omega}$, and equation (3) would reduce to $H_{equal} = \log_2 \Omega$. Since $H_{equal} \geq H_{before}$[14,15], $\Omega \geq 2^{10^{23}}$ states. Yet, after the coin has settled on one side, the uncertainty is only one bit lower because there are half as many microstates: if $H_{before} = log_2(1\Omega)$ and $H_{after} = log_2(\Omega/2)$ then $R = H_{before} - H_{after}$ $= log_2(1\Omega/(\Omega/2)) = 1$ bit.

This assumes, of course, that either result of the coin flip is useful for some function. A coin-flip operation by a molecular machine can be useful if either result helps the survival of the organism that makes the machine. A striking molecular example is the mechanism used by the immune system, where the random joining of gene segments helps to insure the creation of a wide variety of antibodies[10].

However, random choices are not repeatable, so they are not useful to most molecular machines. If a coin flip mechanism were to be used, $H_{before} = log_2(2\Omega)$ but *in the ensemble of all possible after states*, H_{after} also equals $log_2(2\Omega)$, so $R = 0$. No information could be gained in the long run. For example, if the restriction enzyme *Eco*RI did not reliably and repeatably recognize one pattern, GAATTC, the bacterium might die by the destruction of its own genetic material [32]. Likewise, if a DNA polymerase did not reliably insert adenosine opposite every thymidine, many mutations would occur. It is not "simply a matter of putting in the right one" (as we often have a tendency to think); biological systems evolve to avoid mistakes. Macroscopic communications devices must also select one particular state from several possible states. For example, a teletype selects only one character from many incorrect ones because, at any given moment, there is only one correct character to be printed. All others are errors. In both human and biological machines, there is a bias toward one particular state which is preferentially chosen from several possible states.

Even a very energetic penny can gain only one bit of information when it settles down. The following shows that there is a minimum amount of energy that a coin has to give up to specify heads or tails.

Derivation of $\mathcal{E}_{min}$ from the Second Law of Thermodynamics

Instead of a coin, the thermodynamic system we will consider is a single molecular machine. The Second Law of Thermodynamics must apply here, if it is to apply at all[33]. Therefore we may write the Clausius inequality [34,19,35]:

$$dS \geq \frac{dq}{T}. \tag{13}$$

That is, in a small volume that exactly encloses the molecular machine, if a small amount of heat energy dq enters the volume, then the entropy of the molecular machine must increase (dS) by at least $\frac{dq}{T}$, where T is the absolute temperature in K.

Molecular machines operate at one temperature[11], so T is a constant and we may integrate (13) to obtain:

$$\Delta S \geq \frac{q}{T} \tag{14}$$

where q is the total heat entering the volume. By substituting (11) into (14) and rearranging, we obtain a relationship between the information R and the heat q:

$$k_{\mathrm{B}} T \ln(2) \leq \frac{-q}{R} \quad \text{(joules per bit).} \tag{15}$$

The interpretation of this equation is straightforward. There is a minimum amount of heat energy:

$$\mathcal{E}_{min} = k_{\mathrm{B}} T \ln(2) \quad \text{(joules per bit)} \tag{16}$$

that must be dissipated (*negative* q) by a molecular machine in order for it to gain $R = 1$ bit of information. More energy than $\mathcal{E}_{min}$ could be dissipated for each bit gained, but that would be wasteful. This derivation, which consists of definitions and simple rearrangements, shows that (15) and (16) are just restatements of the Second Law under isothermal conditions.

Derivation of $\mathcal{E}_{min}$ from the Capacity of Molecular Machines

The capacity of a molecular machine is given by:

$$C_y = d_{space} \log_2 \left(\frac{P_y}{N_y} + 1 \right) \quad \text{(bits per operation).} \tag{17}$$

[11]. The symbols have the following meanings:

- C_y. The "machine capacity". Closely related to Shannon's channel capacity [29], it is the maximum amount of information which a molecular machine can gain per operation.

- d_{space}. The number of independent parameters needed to define the positions of machine parts. d_{space} cannot be larger than $3n - 6$, where n is the number of atoms in the machine.

- P_y. The "power" or rate at which the machine dissipates energy into the surrounding environment during an operation, in joules per operation.

- N_y. The "noise" or thermal energy which disturbs the machine, in joules.

By dividing the power by the machine capacity at that power we obtain the number of joules that must be dissipated to gain a bit of information[36]:

$$\mathcal{E} \equiv \frac{P_y}{C_y} \quad \text{(joules per bit).} \tag{18}$$

Although decreasing P_y decreases $\mathcal{E}$, the capacity C_y also decreases according to equation (17), so we might incorrectly anticipate that at $P_y = 0$ we would discover that $\mathcal{E}$ would be undefined or zero. However, $\mathcal{E}$ does approach a distinct limit (Fig. 1)[36] which we can find by substituting (17) into (18): ⇐Fig 1

$$\mathcal{E} = \frac{P_y}{d_{space}\log_2\left(\frac{P_y}{N_y}+1\right)} \quad \text{(joules per bit),} \tag{19}$$

and defining $\mathcal{E}_{min}$ as the limit as $P_y \to 0$ (using l'Hôpital's rule[37]):

$$\mathcal{E}_{min} \equiv \lim_{P_y \to 0} \mathcal{E} = \frac{N_y \ln(2)}{d_{space}} \quad \text{(joules per bit).} \tag{20}$$

The thermal noise disturbing a molecular machine is:

$$N_y = d_{space} k_{\mathrm{B}} T \quad \text{(joules)} \tag{21}$$

[11] so substituting (21) into (20) gives us

$$\mathcal{E}_{min} = k_{\mathrm{B}} T \ln(2) \quad \text{(joules per bit)} \tag{22}$$

which is equation (16) again. The value of d_{space}, which is not easy to determine, conveniently drops out of the equation.

This derivation was first recognized by Pierce and Cutler[38,36]. Because it produces the same result as equation (16), the derivation shows that the machine capacity (equation (17)) is closely related to the Second Law of Thermodynamics under isothermal conditions.

Although the present paper was written using the equations for a simple molecular machine, one also obtains equation (22) for both the Shannon receiver[38,36] and for the general molecular receiver[11] because the factors of d_{space} and W cancel between the capacity and noise formulas in each case. (See Table 1 in[11].) So Shannon's channel capacity is, surprisingly, also related to the "isothermal" Second Law of Thermodynamics.

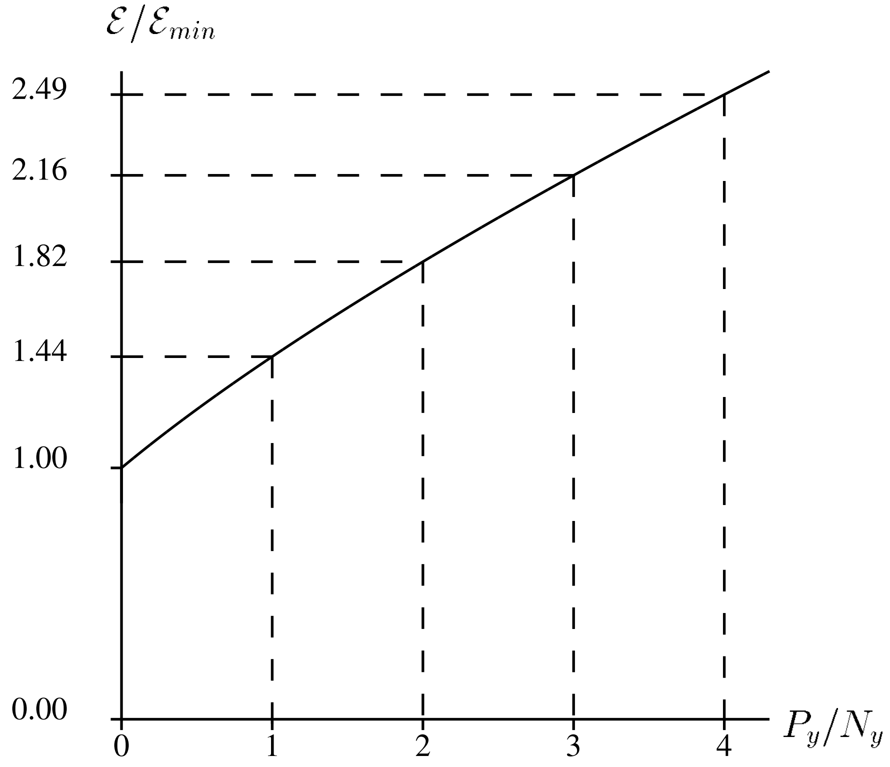

Figure 1: The lower bound on $\mathcal{E}$ is $\mathcal{E}_{min}$.

Logical Operations and Computation by Molecular Machines All molecular machines perform logical operations. For example, if one strand of DNA contains 5′ TAC 3′, then a complete and correct hybridization operation requires that the complementary strand contain 3′ A $\mathcal{AND}$ T $\mathcal{AND}$ G 5′. Likewise the restriction enzyme *Eco*RI cuts DNA only with the pattern 5′ G $\mathcal{AND}$ A $\mathcal{AND}$ A $\mathcal{AND}$ T $\mathcal{AND}$ T $\mathcal{AND}$ C 3′ while other restriction enzymes will bind to only one DNA pattern $\mathcal{OR}$ another[39,40], and the *lac* repressor protein will bind the operator only if it is $\mathcal{NOT}$ also binding an inducer[10]. Any logical function, including $\mathcal{OR}$, addition, and the other algebraic operations, can be constructed entirely from $\mathcal{AND}$ and $\mathcal{NOT}$[41,42,43]. According to the channel capacity theorem[29,11] even operations performed by individual molecules can be precise and almost error free.

Bennett and Landauer[44,45] have proposed that it is not necessary to dissipate energy in order to perform computations. We can show that this is correct by using examples from molecular biology. For example, *Eco*RI effectively performs Boolean logic every time it binds to DNA. Since any computation can be reduced to Boolean operations, *Eco*RI will do arbitrarily large amounts of "computation" when it is non-specifically bound to a DNA that does not contain its binding sites. (The result of the computation in this case is FALSE since some of the bases do not match the required pattern.) However, *Eco*RI must dissipate energy in order to bind at GAATTC. Therefore *each completed operation ("output") performed by a molecular machine in the presence of thermal noise must be accompanied by a dissipation of energy*, according to the Second Law of Thermodynamics, equation (15). That is, although computation does not have an energetic bound, output does. This distinction was recognized by Feynman[46]. Recognizing that output costs at least $\mathcal{E}_{min}$ joules per bit while computation itself is energetically unlimited resolves a long standing dispute[47,48,49,50,51,52,53,44,45,54].

Discussion The derivation of $\mathcal{E}_{min}$ from the Second Law of Thermodynamics is almost certainly the one that von Neumann gave during his lectures at the University of Illinois in 1949[55]. Ironically, his exact words were lost because of noise in a bad tape recording (equation (21)!), and he died before he could complete his book. $\mathcal{E}_{min}$ has been derived in other ways[21,56,57,58] that do not demonstrate its generality.

Equations (15), (16) and (20) are "nothing more than" restatements of the Second Law (equation (13))[4,2]. The derivation holds not only for the machine capacity, but also for Shannon's channel capacity[29] and the general molecular receivers described in[11]. Thus all three theories described in the appendix of[11] give the same value for $\mathcal{E}_{min}$. It is surprising that the close relationship between the Second Law and the channel capacity is not well recognized, since the channel capacity formula has been known since 1949.

In the general molecular receiver theory[11], there are two ways for the power to approach zero to attain the limit $\mathcal{E}_{min}$ when the temperature is held constant. Since

$$P_z = \frac{-q}{t} \qquad \text{(joules per second)}, \tag{23}$$

one of these is to decrease the amount of energy dissipated, $-q \to 0$, while the other is to increase the amount of time, t, that the machine takes to perform its decoding operation. Thus taking the limit as $P_z \to 0$ corresponds to taking the limit as $t \to \infty$ when the energy dissipation $-q$ is held constant. Since this limit produces the isothermal Second Law, and since we all have been taught that the equality in the Second Law only holds for "reversible" machines, we have here a particularly neat way to see the Second Law as the limit of extremely slow "reversible" operations (equation (20)). The same argument holds for Shannon's theory. In contrast, simple molecular machines[11] cannot take advantage of long time periods and $P_y = -q$, so only $P_y \to 0$ is relevant.

Because Shannon's channel capacity theorem applies to formula (18)[29,11], we can see that

> $\frac{1}{\mathcal{E}_{min}} = \frac{1}{k_B T \ln(2)}$ *bits gained per joule dissipated is a precise upper bound on what can be done by a molecular machine.*

The word "precise" means that so long as the bound is not exceeded, the error rate may be made as small as desired.

Another consequence of the channel capacity theorem is that even single molecules can perform precise Boolean logic if they do not exceed the machine capacity. This suggests that fast and accurate molecular computers are possible [59,60,61,62,63,64,65,66,67,68,69] and that these may operate close to $\mathcal{E}_{min}$. Although we do not know how to design them yet, computers built from proteins are well within our present construction capabilities[70,71,72,73,74,75,76].

I thank Herb Schneider and John Spouge for enormously fun and useful discussions, John Skilling for a supportive letter, Peter Basser, Peter Lemkin, Sarah Lesher, Joe Mack, Jake Maizel, Peter Rogan, Denise Rubens and Morton Schultz for critically reading the manuscript and discussing these ideas. I also thank Gary Stormo for pointing out that base pairing requires a 1 in 4 choice, and Larry Gold for supporting the preliminary stages of this project under NIH grant GM28755.

References

1. Maxwell, J. C. (1904). *Theory of Heat*. Longmans, Green and Co., London.

2. Ehrenberg, W. (1967). Maxwell's demon. *Sci. Am.* **217** (5), 103–110.

3. Rothstein, J. (1951). Information, measurement, and quantum mechanics. *Science,* **114**, 171–175.

4. Jaynes, E. T. (1988). The evolution of Carnot's principle. In *Maximum-Entropy and Bayesian Methods in Science and Engineering*, (Erickson, G. J. & Smith, C. R., eds), vol. 1, pp. 267–281, Kluwer Academic Publishers, Dordrecht, The Netherlands. http://bayes.wustl.edu/etj/articles/ccarnot.ps.gz http://bayes.wustl.edu/etj/articles/ccarnot.pdf.

5. Szilard, L. (1929). Uber die entropieverminderung in einem thermodynamischen system bei eingriffen intelligenter wesen. *Z. Phys.* **53**, 840–856.

6. Szilard, L. (1964). On the decrease of entropy in a thermodynamic system by the intervention of intelligent beings. *Behavioral Science,* **9**, 301–310.

7. Feld, B. T. & Szilard, G. W. (1972). *The Collected Works of Leo Szilard: Scientific Papers*. The MIT Press, London.

8. Brillouin, L. (1951). Maxwell's daemon cannot operate: information and entropy. I. *J. of Applied Physics,* **22**, 334–337.

9. Brillouin, L. (1951). Physical entropy and information. II. *J. of Applied Physics,* **22**, 338–343.

10. Watson, J. D., Hopkins, N. H., Roberts, J. W., Steitz, J. A. & Weiner, A. M. (1987). *Molecular Biology of the Gene*. fourth edition, The Benjamin/Cummings Publishing Co., Inc., Menlo Park, California.

11. Schneider, T. D. (1991). Theory of molecular machines. I. Channel capacity of molecular machines. *J. Theor. Biol.* **148**, 83–123. http://www.lecb.ncifcrf.gov/~toms/paper/ccmm/.

12. Britten, R. J. & Kohne, D. E. (1968). Repeated sequences in DNA. *Science,* **161**, 529–540.

13. Callen, H. B. (1985). *Thermodynamics and an Introduction to Thermostatistics*. second edition, John Wiley & Sons, Ltd., N. Y.

14. Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Tech. J.* **27**, 379–423, 623–656. http://cm.bell-labs.com/cm/ms/what/shannonday/paper.html.

15. Shannon, C. E. & Weaver, W. (1949). *The Mathematical Theory of Communication*. University of Illinois Press, Urbana.

16. Pierce, J. R. (1980). *An Introduction to Information Theory: Symbols, Signals and Noise*. second edition, Dover Publications, Inc., New York.

17. Bharath, R. (1987). Information theory. *Byte,* **12** (14), 291–298.

18. Waldram, J. R. (1985). *The Theory of Thermodynamics.* Cambridge University Press, Cambridge.

19. Weast, R. C., Astle, M. J. & Beyer, W. H. (1988). *CRC Handbook of Chemistry and Physics*. CRC Press, Inc., Boca Raton, Florida.

20. von Neumann, J. (1963). Probabilistic logics and the synthesis of reliable organisms from unreliable components. In *Collected works*, (Taub, A. H., ed.), vol. 5, pp. 341–342, The MacMillan Company, New York.

21. Brillouin, L. (1962). *Science and Information Theory*. second edition, Academic Press, Inc., New York.

22. Tribus, M. & McIrvine, E. C. (1971). Energy and information. *Sci. Am.* **225** (3), 179–188. (Note: the table of contents in this volume incorrectly lists this as volume **224**).

23. Schneider, T. D., Stormo, G. D., Gold, L. & Ehrenfeucht, A. (1986). Information content of binding sites on nucleotide sequences. *J. Mol. Biol.* **188**, 415–431. http://www.lecb.ncifcrf.gov/~toms/paper/schneider1986/.

24. Popper, K. R. (1967). Time's arrow and feeding on negentropy. *Nature,* **213**, 320.

25. Popper, K. R. (1967). Structural information and the arrow of time. *Nature,* **214**, 322.

26. Wilson, J. A. (1968). Increasing entropy of biological systems. Entropy, not negentropy. *Nature,* **219**, 534–536.

27. Ryan, J. P. F. J. (1972). Information, entropy and various systems. *J. Theor. Biol.* **36**, 139–146.

28. Ryan, J. P. (1975). Aspects of the Clausius-Shannon identity: emphasis on the components of transitive information in linear, branched and composite physical systems. *Bull. of Math. Biol.* **37**, 223–254.

29. Shannon, C. E. (1949). Communication in the presence of noise. *Proc. IRE,* **37**, 10–21.

30. Hobson, A. (1971). *Concepts in Statistical Mechanics*. Gordon and Breach Science Publishers, New York.

31. Schneider, T. D. & Stormo, G. D. (1989). Excess information at bacteriophage T7 genomic promoters detected by a random cloning technique. *Nucleic Acids Res.* **17**, 659–674.

32. Heitman, J., Zinder, N. D. & Model, P. (1989). Repair of the *Escherichia coli* chromosome after *in vivo* scission by the *Eco*RI endonuclease. *Proc. Natl. Acad. Sci. USA,* **86**, 2281–2285.

33. McClare, C. W. F. (1971). Chemical machines, Maxwell's demon and living organisms. *J. Theor. Biol.* **30**, 1–34.

34. Castellan, G. W. (1971). *Physical Chemistry*. second edition, Addison-Wesley Publishing Company, Reading, Mass.

35. Atkins, P. W. (1984). *The Second Law*. W. H. Freeman and Co., N. Y.

36. Raisbeck, G. (1963). *Information Theory*. Massachusetts Institute of Technology, Cambridge, Massachusetts.

37. Thomas, G. B. (1968). *Calculus and Analytic Geometry*. fourth edition, Addison-Wesley, Reading, Mass.

38. Pierce, J. R. & Cutler, C. C. (1959). Interplanetary communications. In *Advances in Space Science, Vol. 1*, (Ordway, III, F. I., ed.), pp. 55–109, Academic Press, Inc., N. Y.

39. Smith, H. O. (1979). Nucleotide sequence specificity of restriction endonucleases. *Science,* **205**, 455–462.

40. Roberts, R. J. (1989). Restriction enzymes and their isochizomers. *Nucl. Acids Res., Supplement,* **17**, r347–r387.

41. Wait, J. V. (1967). Symbolic logic and practical applications. In *Digital Computer User's Handbook*, (Klerer, M. & Korn, G. A., eds), pp. 4–3 to 4–28 show that NAND is sufficient, McGraw-Hill Book Company, Inc., N. Y.

42. Gersting, J. L. (1986). *Mathematical structures for computer science*. second edition, W. H. Freeman and Co., New York.

43. Schilling, D. L., Belove, C., Apelewicz, T. & Saccardi, R. J. (1989). *Electronic circuits, discrete and integrated*. third edition, McGraw-Hill, New York.

44. Bennett, C. H. (1987). Demons, engines and the Second Law. *Sci. Am.* **257** (5), 108–116.

45. Landauer, R. (1988). Dissipation and noise immunity in computation and communication. *Nature,* **335**, 779–784.

46. Feynman, R. P. (1987). Tiny computers obeying quantum mechanical laws. In *New Directions in Physics: The Los Alamos 40th Anniversary Volume*, (Metropolis, N., Kerr, D. M. & Rota, G., eds), pp. 7–25, Academic Press, Inc., Boston.

47. Bennett, C. H. (1973). Logical reversibility of computation. *IBM J. Res. Develop.* **17**, 525–532.

48. Bennett, C. H. (1982). The thermodynamics of computation - a review. *Int. J. Theor. Phys.* **21**, 905–940.

49. Robinson, A. L. (1984). Computing without dissipating energy. *Science,* **223**, 1164–1166.

50. Porod, W., Grondin, R. O., Ferry, D. K. & Porod, G. (1984). Dissipation in computation. *Physical Review Letters,* **52**, 232–235.

51. Bennett, C. H. & Landauer, R. (1985). The fundamental physical limits of computation. *Sci. Am.* **253** (1), 48–56.

52. Mayer, D. F., Mauldin, J. H., Bennett, C. H. & Landauer, R. (1985). Letters. *Sci. Am.* **253** (4), 6–9.

53. Hastings, H. M. & Waner, S. (1985). Low dissipation computing in biological systems. *BioSystems,* **17**, 241–244.

54. Keyes, R. W. (1989). Making light work of logic. *Nature,* **340**, 19.

55. von Neumann, J. (1966). Fourth University of Illinois lecture. In *Theory of Self-Reproducing Automata*, (Burks, A. W., ed.), pp. 66–67, University of Illinois Press, Urbana.

56. Landauer, R. (1961). Irreversibility and heat generation in the computing process. *IBM J. Res. Dev.* **5**, 183–191.

57. Keyes, R. W. & Landauer, R. (1970). Minimal energy dissipation in logic. *IBM J. Res. Dev.* **14**, 152–157.

58. Keyes, R. W. (1970). Power dissipation in information processing. *Science,* **168**, 796–801.

59. Feynman, R. P. (1961). There's plenty of room at the bottom. In *Miniaturization*, (Gilbert, H. D., ed.), pp. 282–296, Reinhold Publishing Corporation, New York.

60. Drexler, K. E. (1981). Molecular engineering: an approach to the development of general capabilities for molecular manipulation. *Proc. Natl. Acad. Sci. USA,* **78**, 5275–5278.

61. Carter, F. L. (1984). The molecular device computer: point of departure for large scale cellular automata. *Physica D,* **10**, 175–194.

62. Haddon, R. C. & Lamola, A. A. (1985). The molecular electronic device and the biochip computer: present status. *Proc. Natl. Acad. Sci. USA,* **82**, 1874–1878.

63. Conrad, M. (1985). On design principles for a molecular computer. *Comm. ACM,* **28** (5), 464–480.

64. Conrad, M. (1986). The lure of molecular computing. *IEEE Spectrum,* **23** (10), 55–60.

65. Drexler, K. E. (1986). *Engines of Creation*. Anchor Press, Garden City, New York.

66. Arrhenius, T. S., Blanchard-Desce, M., Dvolaitzky, M., Lehn, J. & Malthete, J. (1986). Molecular devices: caroviologens as an approach to molecular wires—synthesis and incorporation into vesicle membranes. *Proc. Natl. Acad. Sci. USA,* **83**, 5355–5359.

67. Hong, F. T. (1986). The bacteriorhodopsin model membrane system as a prototype molecular computing element. *BioSystems,* **19**, 223–236.

68. Hopfield, J. J., Onuchic, J. N. & Beratan, D. N. (1988). A molecular shift register based on electron transfer. *Science,* **241**, 817–820.

69. Eigler, D. M. & Schweizer, E. K. (1990). Positioning single atoms with a scanning tunnelling microscope. *Nature,* **344**, 524–526.

70. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982). *Molecular Cloning, A Laboratory Manual*. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

71. Beaucage, S. L. & Caruthers, M. H. (1981). Deoxynucleoside phosphoramidites - a new class of key intermediates for deoxypolynucleotide synthesis. *Tetrahedron Letters,* **22**, 1859–1862.

72. Pabo, C. (1983). Molecular technology: designing proteins and peptides. *Nature,* **301**, 200.

73. Ulmer, K. M. (1983). Protein engineering. *Science,* **219**, 666–671.

74. Rastetter, W. H. (1983). Enzyme engineering. *Appl. Biochem. and Biotech.* **8**, 423–436.

75. Wetzel, R. (1986). What is protein engineering? *Protein Engineering,* **1**, 3–5.

76. Lesk, A. M. (1988). Introduction: protein engineering. *BioEssays,* **8**, 51–52.