

Entropy of Living versus Non-Living Systems

Isaac C. Sanchez

Department of Chemical Engineering, University of Texas Austin, Austin, USA

E-mail: sanchez@che.utexas.edu

Received January 19, 2011; revised February 17, 2011; accepted March 7, 2011

Abstract

Using a careful thermodynamic analysis of unfertilized and fertilized eggs as a paradigm, it is argued that neither classical nor statistical thermodynamics is able to adequately describe living systems. To rescue thermodynamics from this dilemma, the definition of entropy for a living system must expand to acknowledge the latent genetic information encoded in DNA and RNA. As a working supposition, it is proposed that gradual unfolding (expression) of genetic information contributes a negative entropy flow into a living organism that alleviates apparent thermodynamic inconsistencies. It is estimated that each coding codon in DNA intrinsically carries about -3 k in negative entropy. Even prior to the discovery of DNA and the genetic code, negative entropy flow in living systems was first proposed by Erwin Schrödinger in 1944.

Keywords: Negative Entropy, DNA, Genetic Code, Codon, Second Law of Thermodynamics

1. Introduction

Applying thermodynamics to living systems often frustrates intuition and challenges comprehension [1,2]. The conundrum is exemplified by a careful thermodynamic analysis of a system in which unfertilized and fertilized eggs are placed into a large adiabatic incubator. For each egg the requisite global requirements of the second law, *viz.*, a positive entropy change, are satisfied. But it is also shown that chemical reactions in the fertilized egg also *increase* its entropy. This result belies the development of an embryo within the fertilized egg. To rescue thermodynamics from this dilemma, an entropy definition for a living system must acknowledge the *latent* genetic information encoded in DNA and RNA. As a working supposition, it is proposed that gradual unfolding (expression) of genetic information contributes a negative entropy flow into a living organism that alleviates apparent thermodynamic inconsistencies. An admittedly crude attempt to calculate this negative entropy contribution is also outlined. This radical shift in paradigm may prove useful to model aging as well as other biological processes.

2. Egg Conundrum

The thought experiment involves placing an unfertilized and fertilized egg into a large adiabatic chamber that contains heated air at atmospheric pressure, at a fixed

temperature, and with a fixed relative humidity. The volume of air in the adiabatic chamber is large enough so that heat transfers to or from the eggs don't change the air temperature appreciably (a large heat reservoir). This adiabatic system is isolated from the rest of the universe and energy is conserved within (first law).

Unfertilized egg: The egg at room temperature (T_1) is placed into the adiabatic chamber at the incubation temperature (T_2). It is held there until it comes to thermal equilibrium with the heated air. As the egg heats up from T_1 to the incubation temperature T_2 it absorbs an amount of heat, $Q_{12} > 0$. Applying the first law (all symbols have their customary meanings):

$$\Delta U_{egg} = Q_{12} = C_v \Delta T; \Delta U_{air} = -Q_{12} \quad (1)$$

$$\begin{aligned} \Delta S_{egg} &= C_v \ln(T_2/T_1) \\ &= Q_{12} \ln(T_2/T_1)/\Delta T > 0, \Delta S_{air} = -Q_{12}/T_2 < 0 \end{aligned} \quad (2)$$

$$\begin{aligned} \Delta S_{egg} + \Delta S_{air} &= -\frac{Q_{12}}{T_2} \left[\frac{\ln(1 - \Delta T/T_2)}{\Delta T/T_2} \right] - \frac{Q_{12}}{T_2} \\ &= \frac{Q_{12}}{T_2} \left[1 + \frac{1}{2} \left(\frac{\Delta T}{T_2} \right) + \dots \right] - \frac{Q_{12}}{T_2} = \frac{1}{2} \frac{Q_{12}}{T_2} \left(\frac{\Delta T}{T_2} \right) + \dots > 0 \end{aligned} \quad (3)$$

and the second law is satisfied.

Fertilized egg: Heating the fertilized egg to the incubation temperature will increase its entropy as above, but there are some additional changes. Heating to the incu-

bation temperature triggers irreversible chemical reactions within the fertilized egg. The sum of these reactions will generate or absorb heat. Let Q_{rx} be the net heat generated or absorbed by the egg caused by chemical reactions. So in addition to the above entropy change, the fertilized egg will experience an additional entropy change, ΔS_{rx} . Let Q_{air} be the heat transfer from the air at the incubation temperature T_2 from or to the egg. To maintain the egg at the incubation temperature, any excess heat generated by chemical reactions is dumped to the reservoir and any heat absorbed by the chemical reactions must be supplied by the reservoir. Thus, for either case energy is conserved and $Q_{rx} + Q_{air} = 0$. Phase transformations with their associated heats are also gathered into Q_{rx} .

Now the second law requires

$$\Delta S_{rx} + \Delta S_{air} > 0 \quad \text{or} \quad \Delta S_{rx} + \frac{Q_{air}}{T_2} > 0 \quad (4)$$

Experience suggests that embryonic growth within a fertilized egg requires a net absorption of heat during incubation; *i.e.*, the irreversible chemical reactions within the egg require absorption of heat energy from the surroundings ($Q_{air} < 0$). The second law then implies that $\Delta S_{rx} > 0$. Gas transport across the eggshell is ignored, but this minor effect can be accounted for if necessary.

So if chemical reactions cause the egg to absorb heat, the entropy of the egg must increase. Of course, this entropy increase completely disagrees with what is going on inside the egg, namely the growth of an embryo. The alternative is that the chemical reactions generate heat energy leading to a decrease in the egg entropy. However, this scenario can be ruled out by application of *Le Chatelier's principle*. Raising the egg temperature to the incubation temperature to promote chemical reactions requires the consumption of heat energy. If these reactions were energy producing, they would slow down with an increase in temperature. Moreover, it is also well-known that to sustain life, cells require a continuous input of energy.

As a parenthetical remark, if a black box filled with chemicals was placed inside the adiabatic incubator and the rise in temperature triggered a cascade of chemical reactions, the thermodynamic analysis of the black box would be identical to that of the fertilized egg. The big difference, of course, is no life form would have been created.

3. Thermodynamic Rescue

To rescue thermodynamics from this apparent dilemma, it is proposed that the definition of entropy expand to properly include living systems. The need for this expansion

can be illustrated by comparing the entropy of a perfect crystal at 0 Kelvin with the entropy of a strand of DNA at room temperature. According to the third law, the entropy of the perfect crystal is zero. From the perspective that statistical entropy measures the number of microstates available to the system, the perfect crystal has lower entropy than DNA. But from the perspective of information content, DNA has lower entropy than the perfect crystal. Boltzmann's definition of statistical entropy does not describe nor capture the informational content encoded in DNA or RNA. However, this informational entropy, which is negative relative to a perfect crystal, is *latent* and not always expressed. As long as the genetic information remains unexpressed, as in the incubation of an unfertilized egg, standard thermodynamics offers an accurate account of entropy changes.

Another *gedanken* experiment that illustrates the inability of standard thermodynamics to capture encoded information is the following: Imagine synthesizing a strand of RNA of fixed length that produces a random sequence of nucleotides. Carrying out the reaction in a calorimeter, the heat of reaction is measured along with its temperature dependence. This would allow the calculation of the entropy, the enthalpy, and free energy of reaction. Now do this many times, each time synthesizing a different random sequence of RNA nucleotides. The measured thermodynamic properties would not differ very much from one random RNA strand to another because of the chemical similarity of the nucleotides. Now if one of these RNA's by chance matched the RNA transcribed from the DNA of a living organism, it would be carrying coded information for the production of a specific functional protein. However, thermodynamic measurements would never detect this unique and informational containing RNA strand.

Our conclusion is that the essential macromolecules of living systems, DNA and RNA, appear to require an additional entropy descriptor that lies outside the scope of classical or statistical thermodynamics. When the encoded information is expressed, a negative contribution contributes to the classical entropy of the living system. So in a living system two principles are in operation: the usual one for non-living systems that always proceeds with an increase in *global* entropy and a second one, superimposed on the first that proceeds with a decrease in *local* entropy of the living organism. The latter is unaccounted for in classical or statistical thermodynamics. Thus, the *local* entropy changes in a living organism are given by

$$\Delta S_{living} = \Delta S_{class} + \Delta S_{gene} < 0 \quad (5)$$

where the classical entropy ΔS_{class} will be positive, but the entropy contribution from expressed genetic information (ΔS_{gene}) is always negative and will overwhelm the

classical contribution in a living organism. Even before the discovery of DNA and the genetic code, Schrödinger [2] had already come to a similar conclusion:

“Thus, a living organism continually increases its entropy—or, as you may say, produces positive entropy—and thus tends to approach the dangerous state of maximum entropy, which is death. It can only keep aloof from it, i.e. alive, by continually drawing from its environment of negative entropy ... what an organism feeds upon is negative entropy.”

A living system is a non-equilibrium system in which the entropy is changing. Equation (5) is very similar to the well-known equation of non-equilibrium thermodynamics that describes entropy flow into a system where irreversible processes are taking place [3]. The only difference is that an additional term would be added to reflect the negative entropy flow from the unfolding of genetic information.

4. Incomplete Quantification of Negative Entropy

The challenge is to quantify this non-classical entropy. Information theory has already made some progress in measuring the information content in DNA [4]. Another statistical approach is as follows: the genes in DNA all have a specified sequence of codons. Let p^k be the probability that the unique sequence that defines the k^{th} gene would randomly occur; this probability is very small and equals $e^{-g_k n_k}$ where g_k is a positive constant and n_k is the number of codons that code the k^{th} gene. The probability P_N assigned to the m genes containing a total of N codons in a strand of DNA is given by:

$$P_N = \prod_{k=1}^m p_k = \exp \left[-\sum_{k=1}^m g_k n_k \right] \\ = \exp \left[-\frac{\sum_{k=1}^m g_k n_k}{\sum_{k=1}^m n_k} \sum_{k=1}^m n_k \right] \equiv \exp[-\bar{g}N] \quad (6)$$

where $N = \sum n_k$ is the total number of coding codons in the DNA strand and \bar{g} is the average value of the degeneracy factors, g_k , that can always be evaluated for any specified sequence. Since there are only 20 coded amino acids, some of the 64 codons are degenerate. For example, the sequence ATGGAAGCTCGCAAA codes for the amino acid sequence methionine-glutamic acid-alanine-arginine-lysine. The probability of this five-fold sequence arising randomly from the 64 possible codons is $(1/64)(2/64)(4/64)(6/64)(2/64) = 96/64^5 = e^{-16.2} = e^{-3.24(5)}$. Thus, $g_k = 3.24$ in this example.

To calculate the entropy of DNA, the generalized

Gibbs entropy is used [5]

$$S = -k \sum_l P_l \ln P_l \quad (7)$$

where P_l is the probability of the l_m state and the sum is over all possible states of the system. In this application, P_l is the probability of a unique linear sequence of N codons and the sum is over all possible sequences. Let $\Omega = \sum_l 1$ be the total number of ways that N codons can be linearly arranged in DNA. Because many of the 64 codons are degenerate, with an average degeneracy of 3, it is estimated that

$$\Omega = \left(\frac{64}{3} \right) = e^{3.06N} \quad (8)$$

Thus, on average each unique arrangement of N codons has a probability $P_l = 1/\Omega$ and

$$S = -k \sum_l \left(\frac{1}{\Omega} \right) \ln \left(\frac{1}{\Omega} \right) = k \ln \Omega = 3.06kN \quad (9)$$

This is the total statistical entropy of N codons in DNA. So relative to the total ensemble of configurations, each codon of a non-random and functional DNA strand contributes a relative negative entropy of

$$S_{\text{codon}} = \ln P_l / N = -3.06k \quad (10)$$

Thus, each coding codon in a DNA or RNA strand intrinsically carries about $-3k$ in negative entropy. A gene of m codons would have a “negative gene entropy” of $-3mk$.

An additional rationale for negative entropy is that when DNA replicates, it does so perfectly almost every time. Now let's imagine a life form where DNA replicated imperfectly. If the random duplications are non-functional, and an exact copy is required, the probability of producing a functional copy is P_N . Comparing the entropy change of this hypothetical random DNA replication with the real one, we see that the real DNA replicates with lower entropy of $\ln P_N$.

For the sake of comparison, a water molecule suffers an entropy loss of $1.2k$ during freezing, which is also the same entropy loss during the fluid-solid transition of hard spheres [6], and thus comparable to the $-3k/\text{codon}$ or $-k$ per coding nucleotide. A better perspective is gained by looking at the set of chromosomes in a cell that makes up its genome; it is estimated that the human genome has approximately 3 billion base pairs or 1 billion codons/strand of DNA arranged into 46 chromosomes [7]. Since there are about 10^{14} cells in the human body, there is approximately 1 mole of coding codons in a human. This is a surprisingly small number and our estimate of the negative entropy associated with DNA is most likely a gross underestimate. Less than 2% of the base pairs in DNA are involved in gene coding [8]. It is still not clear that the remaining 98%, the so-called “junk

DNA,” does not contain considerable information content [8,9]. But even taking the junk DNA into account will not likely resolve the issue.

The larger problem is that gene expression and resulting cellular activities depend on more than the genetic code. The genetic code is controlled by gene regulation [10] and executed by a poorly understood cooperation of sophisticated molecular machines, such as ribosomes, DNA and RNA polymerases, different functional types of RNA, DNA enhancers and silencers, molecular motors, enzymes, proteins, among others. These molecular machines work cooperatively to express the coded information in DNA with great precision. Quantifying the negative entropy associated with the “orchestration” necessary for the execution of the genetic code remains a formidable challenge for the future.

However, a bound can be placed on the minimum amount of negative entropy required to sustain growth during incubation of a chick embryo. Let \dot{Q}_{air} represent the heat flow per unit time from the heated incubator air to a fertilized egg. From inequality (4) and Equation (5), we have,

$$|\Delta\dot{S}_{gene}| \gg \frac{\dot{Q}_{air}}{T_2} \quad (11)$$

i.e., the magnitude of the negative entropy flow must far exceed the classical positive entropy increase suffered by the absorption of heat energy by the egg. To the author’s knowledge, the requisite calorimetric measurements have not been made to measure the heat uptake by a fertilized egg in incubation.

5. Aging Model

One possible application of this paradigm shift is in the aging of a life form. For example, as a human develops, from embryo to infant to child to adult, this development could be modeled as a competition between negative and positive entropy flows. Since DNA replicates, negative entropy increases with age, but seems to reach a zenith early in adulthood and a nadir when death is caused by “natural causes.” Positive entropy flow from normal degradative processes, such as destruction of DNA by free radicals, mutations of DNA during replication, poor eating habits, disease, and many other factors, slow the unfolding of negative gene entropy. In this regard, some questions are: if all natural degradative processes could be removed, would the flow of negative gene entropy continue uninterrupted at some optimum level? What is that optimum level? And can it be measured or quantitatively described?

6. Conclusions

Our conclusion is that the essential macromolecules of

living systems, DNA and RNA, appear to require an additional entropy descriptor that lies outside the scope of classical or statistical thermodynamics. Standard thermodynamics is unable to capture the encoded information contained in the DNA and RNA of living systems. This has been illustrated by a careful thermodynamic analysis of fertilized and unfertilized eggs. The apparent dilemma can be resolved by redefining the entropy of a living system to acknowledge the informational content embedded in RNA and DNA. In a living system two principles are in operation: the usual one for non-living systems that always proceeds with an increase in *global* entropy and a second one, superimposed on the first that proceeds with a decrease in *local* entropy of the living organism.

This shift in paradigm may prove useful to model aging of a living organism. A formidable remaining problem is to quantify the flow of negative entropy.

7. References

- [1] A. Pross, “The Driving Force for Life’s Emergence: Kinetic and Thermodynamic Considerations,” *Journal of Theoretical Biology*, Vol. 220, No. 3, 2003, pp. 393-406. [doi:10.1006/jtbi.2003.3178](https://doi.org/10.1006/jtbi.2003.3178)
- [2] E. Schrödinger, “What is Life? The Physical Aspect of a Living Cell,” Cambridge University Press, Cambridge, 1944.
- [3] P. Glandsdorff and I. Prigogine, “Thermodynamic Theory of Structure, Stability, and Fluctuations,” Wiley-Interscience, New York, 1971.
- [4] R. RomanRoldan, P. BernalGalvan and J. L. Oliver, “Application of Information Theory to DNA Sequence Analysis: A Review,” *Pattern Recognition*, Vol. 29, No. 7, 1996, pp. 1187-1194. [doi:10.1016/0031-3203\(95\)00145-X](https://doi.org/10.1016/0031-3203(95)00145-X)
- [5] R. Kubo, “Statistical Mechanics,” North Holland Publishing, Amsterdam, 1965.
- [6] I. C. Sanchez and J. S. Lee, “On the Asymptotic Properties of a Hard Sphere Fluid,” *Journal of Physical Chemistry B*, Vol. 113, No. 47, 2009, pp. 15572-15580. [doi:10.1021/jp901041b](https://doi.org/10.1021/jp901041b)
- [7] J. C. Venter, *et al.*, “The Sequence of the Human Genome,” *Science*, Vol. 291, No. 5507, 2001, pp. 1304-1351. [doi:10.1126/science.1058040](https://doi.org/10.1126/science.1058040)
- [8] E. Pennisi, “Shining a Light on the Genome’s ‘Dark Matter,’” *Science*, Vol. 330, No. 6011, 2010, p. 1614. [doi:10.1126/science.330.6011.1614](https://doi.org/10.1126/science.330.6011.1614)
- [9] E. Pennisi, “Searching for the Genome’s Second Code,” *Science*, Vol. 306, No. 5696, 2004, pp. 632-635. [doi:10.1126/science.306.5696.632](https://doi.org/10.1126/science.306.5696.632)
- [10] S. T. Kosak and M. Groudine, “Gene Order And Dynamic Domains,” *Science*, Vol. 306, No. 5695, 2004, pp. 644-647. [doi:10.1126/science.11038](https://doi.org/10.1126/science.11038)