



HAL
open science

Information-energy equivalence and the emergence of self-replicating biological systems

Emmanuel Paradis

► **To cite this version:**

Emmanuel Paradis. Information-energy equivalence and the emergence of self-replicating biological systems. *BioSystems*, 2023, 226, pp.104885. 10.1016/j.biosystems.2023.104885 . ird-04892566

HAL Id: ird-04892566

<https://ird.hal.science/ird-04892566v1>

Submitted on 16 Jan 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

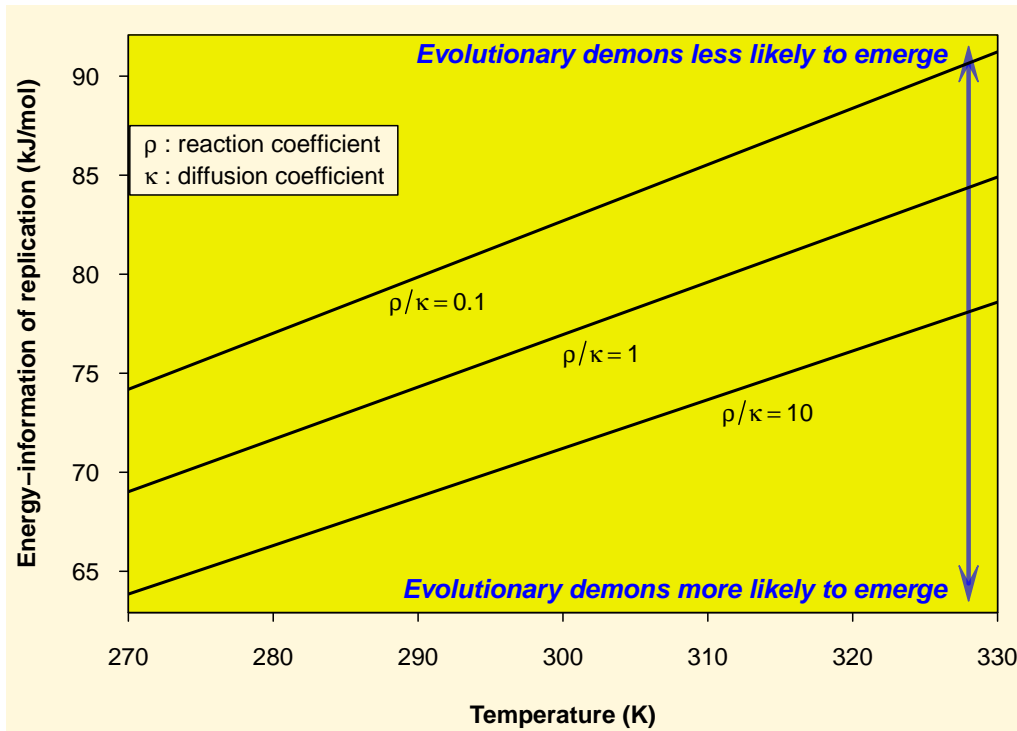
L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Copyright

Graphical Abstract

Information–energy equivalence and the emergence of self-replicating biological systems

Emmanuel Paradis



Highlights

Information–energy equivalence and the emergence of self-replicating biological systems

Emmanuel Paradis

- An equation of the total energy for simple genome replication is derived.
- Decreases in entropy and in Gibbs free energy can be accomplished together.
- Some factors may have favoured the emergence of self-replicating biological systems.
- Short RNA sequences made of G and C may conform to these predictions.

Information–energy equivalence and the emergence of self-replicating biological systems

Emmanuel Paradis

ISEM, IRD, Univ. Montpellier, CNRS, Montpellier, France

ARTICLE INFO

Keywords:

energy
entropy
genome
information
replication

ABSTRACT


Biological processes are characterized by a decrease in entropy in apparent violation of the second law of thermodynamics. Information stored in genomes help to solve this paradox when interpreted under the relationship between information and energy stated by Brillouin in the 1950's. However, the origins of living forms from inanimate matter which have no information storage device remains an open question. In this paper, a theoretical approach is developed on this issue. The replication of a simple entity with a binary genome is assumed to require an information-equivalent energy in addition to the standard activation energy. It is found that, in some conditions, a decrease in entropy can be accomplished together with a decrease in Gibbs free energy. An equation of the total energy for the replication of this entity is derived. Three factors are predicted to lower this energy: a small number of states of the coding sequence, a lower temperature, and a high ratio of the reaction on diffusion coefficients. These factors may have favoured the emergence of evolutionary demons—information storage devices that are able to decrease entropy. It is evaluated that some short, single-stranded RNA sequences made only of G and of C may conform to this model. The consequences of this model and its predictions on the origins of life on Earth and on other planets are discussed.

1 We have already seen that a great deal of information is needed to specify the structure of
2 a complicated nonrandom object. It may be concluded that anything that we would want to
3 call “living” would have a high information content. This apparently simple requirement has
4 far-reaching consequences.

5 *Orgel (1973, p. 192)*

6 **1. Introduction**

7 Brillouin (1949) started his article ‘Life, thermodynamics, and cybernetics’ by asking: “How is it possible to understand life, when the whole world is ruled by such a law as the second principle of thermodynamics, which points toward death and annihilation?” The second law of thermodynamics states that
8 a spontaneous chemical reaction implies an increase of entropy of the system; however, many chemical
9
10

 Emmanuel.Paradis@ird.fr (E. Paradis)
ORCID(s): 0000-0003-3092-2199 (E. Paradis)

11 reactions in living beings imply a decrease in entropy. For instance, the process of DNA duplication results
12 in two DNA molecules identical to themselves thus implying a decrease in entropy. Brillouin's question
13 might be reformulated in the following way: Which physical mechanisms might explain how living beings
14 have evolved from inanimate matter? This question—which is essential to understand the origin of life on
15 Earth and possibly in other natural or artificial systems—remains essentially unsolved. Life on Earth is,
16 as far as we know today, the exception of how matter is organized in the Universe, and physical laws do
17 not to explain this exception (Hawking and Ellis, 1973; Crick, 1981). This is well illustrated by the fact
18 that living beings, as pointed out by Brillouin, appear to violate the second law of thermodynamics which
19 predicts that entropy increases in a closed system. By contrast to what is expected from the second law
20 of thermodynamics, a DNA molecule will replicate identical to itself in a predictable way, thus decreasing
21 entropy. This paradox has attracted a lot of interest from researchers over several decades.

22 Genetic information is stored in genomes which can present a great variety of structures depending on
23 the taxonomic group (see an overview in Paradis, 2020). With a few exceptions of some virus genomes
24 (which can be RNA sequences), genetic information is stored in DNA sequences. Although the physical
25 structure of genomes is important, molecules of DNA are not only simple physical entities: they should be
26 considered as information storage devices. This point is a central feature in the modern concept of life (Vitas
27 and Dobovišek, 2019). A very general definition of an information storage device is: an ordered sequence
28 of discrete units which may take different states. In the case of DNA, there are four such states: {A, C, G,
29 T}. In the case of binary systems, such as digital computers, there are two states: {0, 1}. The link between
30 entropy and information was established by Shannon (1948) who made an almost strict parallel between the
31 classical measure of entropy which considered microstates of physical or chemical entities, and communi-
32 cation exchange between communicating entities. Although the validity of Shannon's framework has been
33 debated (e.g., Kish and Ferry, 2018), some authors, such as Bekenstein (1973), do not see a problem with
34 this parallel. On the other hand, Brillouin (1953) formulated a link between entropy and information which
35 states that if an experiment yields an increase in information concerning a physical system, it must be paid
36 for by a larger increase in the entropy of the system. Brillouin's framework effectively draws links between
37 the concepts of energy, information, and entropy, and helped to solve the paradox of Maxwell's demon by
38 posing that acquiring information necessitates some energy (Szilard, 1964; Earman and Norton, 1998).

39 Information storage, replication, and expression are at the core of life and evolution. These features
40 distinguish living beings from inanimated entities (Vitas and Dobovišek, 2019). Information in living beings
41 must be read to be used or replicated. Information replication is indispensable for survival of living beings.
42 Similarly, the use of stored information, in whatever form, is indispensable for living beings to interact with
43 their environment. It appears therefore that reading information stored in living beings is an indispensable
44 mechanism for their survival. Although the question of replicating and maintaining a genome has been

45 extensively addressed in the literature (e.g., Lynch and Conery, 2003; Elliott and Gregory, 2015), this is
46 generally posed in terms of energy cost with respect to genome sizes rather than in terms of informational
47 contents (Bonnici and Manca, 2016; Manca, 2017).

48 In this paper, I present a theoretical approach linking information theory and the evolution of informa-
49 tion processing at the storage level. The rationale of this framework is the link between information and
50 energy derived from Brillouin's formula on negentropy. The consequences of this framework for genomic
51 evolution and the origins of life are discussed.

52 **2. Theory and models**

53 The approach adopted here is based on a basic model of replication. This is justified in order that
54 the present model and the inferred conclusions are not restricted by the specific physical nature of the
55 information coding structure. Besides, considering that the exact nature of the first living beings on Earth
56 is unknown, an abstract model is probably a good approach here.

57 The present model considers a symbolic representation of a self-replicating entity which is denoted as
58 z . It is assumed that there are several copies in the environment that are not identical. We assume that z is
59 made of two basic elements, the first element is written X and is constant for all copies of the entity, and the
60 second element is either 0 or 1. So we can write $z = \{X0, X1\}$. We assume that the environment contains
61 copies of X , 0, 1 with no limitation.

62 **2.1. Information, energy, and replication**

63 The physical nature of z is not important: what is relevant here is that it can be either $X0$ or $X1$. Unless
64 this entity is perpetual or eternal, it will eventually die or decay. Therefore, to be able to maintain its
65 existence, it has to reproduce through a process of replication. To be effective, this process must be faithful:
66 entities with $z = X0$ must replicate to result in two entities with $z = X0$, and entities with $z = X1$ must do the
67 same to entities with $z = X1$. This faithfulness of the replication process is a prerequisite for life. Another
68 prerequisite is that the physical unit of the information storage must be present in the environment to permit
69 its replication. Therefore, a living entity with a one-bit information storage must acquire at least two bits of
70 information to replicate itself: one bit for the value of z , and one bit from the environment.

The equivalence between energy and information stemming from Brillouin's derivation of the negentropy principle states that acquiring one bit of information requires an energy at least proportional to $k_B \ln 2$ joules where k_B is the Boltzmann constant (Brillouin, 1953, 1959). More specifically, we may write that acquiring one bit of information requires at least ΔE joules and a lower bound for this quantity is given by:

$$\Delta E \geq k_B T \ln 2, \quad (1)$$

where T is the temperature in kelvins. Taking $T = 310.15 \text{ K}$ ($= 37 \text{ }^\circ\text{C}$) and $k_B = 1.380\,649 \times 10^{-23} \text{ J.K}^{-1}$, one finds $\Delta E \geq 2.968\,114 \times 10^{-21} \text{ J}$ for one bit of information. This value is actually meaningful at the microscopic level, that is considering a bit in isolation. At the macroscopic level, we have to consider a large number of information bits and multiply this quantity of energy by Avogadro's number $N = 6.022\,140\,76 \times 10^{23}$, so that we have:

$$\Delta E \geq 1787.44 \text{ J.mol}^{-1}. \quad (2)$$

71 The relevance of these equations for living beings appears when considering that living systems, such as
 72 DNA molecules, must replicate themselves by acquiring information about their surrounding environment.
 73 As mentioned previously, two bits of information must be acquired for the replication of z : one bit to know
 74 the state of z and one bit from the environment. Therefore, the right-hand side term of (2) must be multiplied
 75 by two. At this point, it is still an open question what is the source of the energy required to acquire these
 76 information bits: the system may lower its internal energy level, thus releasing some energy, or energy may
 77 be input from outside the system. The crucial point is that the system requires a source of energy which can
 78 be used as information.

79 Let us consider again our hypothetical living being as above where the information storage is made of a
 80 single bit of information so that z takes the value $X0$ or $X1$. We also assume that the surrounding environ-
 81 ment contains many available copies of 0's and 1's in equal proportions; in other words, this environment
 82 has high (thermodynamic) entropy. In order to replicate itself, our hypothetical simple living being must
 83 acquire information about which molecule is 0 and which molecule is 1. According to the above equations,
 84 this has an energy cost given by ΔE . In a simple system, this cost can be met by decreasing the energy level
 85 during information acquisition. Hydrogen bonds (H-bonds) provide a mechanism that seems suitable for
 86 this. These bonds are ubiquitous in chemical systems (e.g., in water), do not require high energy inputs to be
 87 created (unlike common molecular bonds), and decrease the overall energy level of these systems. There-
 88 fore, we predict that a simple living system with its information stored in a single bit may be self-replicating
 89 if matching 0's and 1's create H-bonds that decrease energy by at least $2 \times 1787.44 = 3574.88 \text{ J.mol}^{-1}$.

This simple model may be extended to nucleic acids (DNA and RNA) since information acquisition in these molecules is based on H-bonds. In the case of DNA sequences, $\ln 2$ in (1) must be replaced by $\ln 4$ because there are four possible bases instead of two. This is equivalent to multiply the previous quantity by

two (since $\ln 4 = 2 \ln 2$). Therefore, we have for DNA sequences:

$$\Delta E \geq 7149.759 \text{ J.mol}^{-1}. \quad (3)$$

90 This prediction can be compared with values from the literature about the energy levels linked to base pair-
91 ing in DNA. Mo (2006) estimated these energies at the following values:

92

$$\Delta E(\text{CG}) = 106.17 \text{ kJ.mol}^{-1} (= 25.4 \text{ kcal.mol}^{-1})$$

93

$$\Delta E(\text{AT}) = 51.83 \text{ kJ.mol}^{-1} (= 12.4 \text{ kcal.mol}^{-1})$$

94 These are in agreement with the above value, and even much larger. We will come back to this point below.

We now attempt to establish the overall energetic needs of replicating our hypothetical living entity. Our model assumes that replication requires a chemical reaction in order to link X with either 0 or 1. The assumption of the need for a chemical reaction for the replication of living beings is likely to be true in many cases (Crick, 1981). We may use Eyring (1935) equation that defines the rate of a chemical reaction denoted as ρ :

$$\rho = \frac{\kappa k_{\text{B}} T}{h} \exp\left(-\frac{\Delta G^{\ddagger}}{RT}\right), \quad (4)$$

where κ is the diffusion coefficient of the reaction, h is the Planck constant, ΔG^{\ddagger} is the activation energy, and R is the gas constant. This activation energy is the quantity of energy that must be provided to the system so that the reaction occurs. In most chemical reactions occurring in living cells, this energy is provided by coupled chemical reactions, such as the hydrolysis of adenosine triphosphate (ATP), and catalyzed by enzymes which lower ΔG^{\ddagger} (Ji, 1974, 2000). In the present simple model of replication, the activation energy must be provided in addition to the energy needed to read information as examined above. By rearranging (4), the activation energy can be written:

$$\Delta G^{\ddagger} = -RT \ln\left(\frac{\rho h}{\kappa k_{\text{B}} T}\right).$$

A chemical reaction is also characterized by the overall change in energy level of the system which is called the Gibbs free energy (Gibbs, 1873), and is defined by:

$$\Delta G^{\circ} = \Delta H^{\circ} - T \Delta S^{\circ},$$

95 where H and S denote the enthalpy and the entropy of the system, respectively, so that ΔH° is the overall
96 change in enthalpy during the reaction, and similarly for ΔS° regarding the entropy of the system (Truhlar

97 et al., 1983). The enthalpy of a system is defined by $H = U + pV$ where U is the internal energy of the
 98 system, p is the pressure, and V is the volume. Assuming that p and V do not change during the reaction,
 99 we can write $\Delta H^\circ = \Delta U^\circ$, the overall change in enthalpy of the system is equal to the overall change in
 100 internal energy of the system and we can write $\Delta G^\circ = \Delta U^\circ - T\Delta S^\circ$.

Both quantities of energy considered above are important to predict whether a chemical reaction is likely to happen: spontaneous reactions have a negative change in free energy ($\Delta G^\circ < 0$); however, if the activation energy is too large such reactions require a similarly large quantity of external energy such as heat or work to happen (Fig. 1). The second law of thermodynamics states that a spontaneous chemical reaction implies an increase in entropy of the system (i.e., $\Delta S^\circ > 0$). However, the replication of z implies by definition a decrease in entropy (as is also the case during replication of living beings). Since a spontaneous reaction is characterized by $\Delta G^\circ < 0$ which implies, in this case, that $\Delta U^\circ < T\Delta S^\circ$. However, we assume here that the second law of thermodynamics is not respected, thus $\Delta S^\circ < 0$ and since T is always strictly positive, we have $-T\Delta S^\circ > 0$. Therefore, to satisfy $\Delta G^\circ < 0$, the following inequalities must hold:

$$\Delta U^\circ < 0$$

$$|\Delta U^\circ| > T|\Delta S^\circ|.$$

101 In other words, the decrease in internal energy of the system must be greater than the decrease in entropy
 102 multiplied by the temperature.

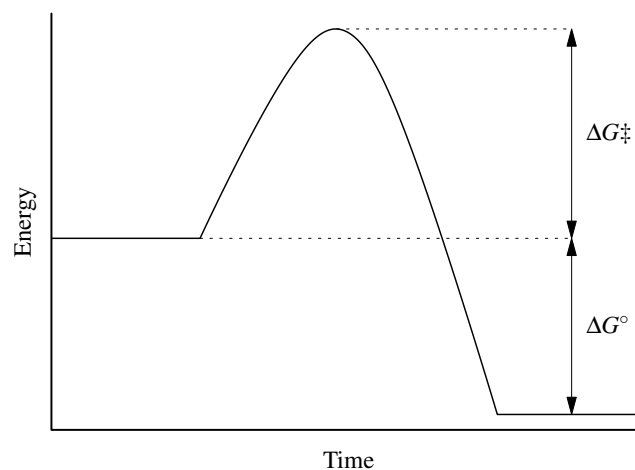


Figure 1: Changes in energy level during a chemical reaction showing the activation energy (ΔG^\ddagger) and the free energy (ΔG°). In this hypothetical example, the reaction produces energy ($\Delta G^\circ < 0$).

We now introduce the energy-information equivalent required for replication, denoted as \mathfrak{E} , and is

defined as the sum of Brillouin's information given by (1) and the activation free energy ΔG^\ddagger :

$$\mathfrak{G} = 2Nk_B T \ln Z - RT \ln \left(\frac{\rho h}{\kappa k_B T} \right),$$

where Z is the number of states of the basic information storage (i.e., $Z = 2$ for the hypothetical entity z , $Z = 4$ for RNA and DNA). The second term of the right-hand side can be split to isolate the two coefficients of the reaction (ρ and κ):

$$\mathfrak{G} = 2Nk_B T \ln Z - RT \ln \frac{\rho}{\kappa} - RT \ln \left(\frac{h}{k_B T} \right).$$

It is a central point of the present hypothesis that the energy required by replication is provided in the above theoretical model by the lowering of internal energy of the system, so this implies:

$$\mathfrak{G} \leq -\Delta U^\circ.$$

In other words, during the chemical reaction involved in replication, the decrease in internal energy must be at least equal to the activation energy of the reaction plus the energy defined by Brillouin's equation, $k_B T \ln 2$. Therefore, any factor that will lower the value of \mathfrak{G} will stimulate chemical reactions leading to self-replication. Let us define three coefficients A , B , and C that depend only on T and on the constants previously defined (k_B , h , and R):

$$A = 2Nk_B T$$

$$B = RT$$

$$C = -RT \ln \left(\frac{h}{k_B T} \right).$$

103 Since $h = 6.626\,070\,15 \times 10^{-34} \text{ J}\cdot\text{Hz}^{-1}$, then $h/k_B \approx 5 \times 10^{-11}$, and thus C will be greater than zero for all re-
 104 alistic values of T . Additionally, A and B are also strictly positive ($R = 8.314\,462\,618\,153\,24 \text{ J}\cdot\text{mol}^{-1}\cdot\text{K}^{-1}$).
 105 The energy-information equivalent of replication is now written as:

$$\mathfrak{G} = A \ln Z - B \ln \frac{\rho}{\kappa} + C.$$

106 It appears that three factors are likely to influence the value of \mathfrak{G} . First, the lower the value of Z , the
 107 lower the value of \mathfrak{G} (since $Z \geq 2$, then $A \ln Z > 0$). Second, the ratio ρ/κ will always be positive (both
 108 parameters are always positive) so that the crucial point is whether this ratio is less than, equal to, or greater

109 than one:

$$\frac{\rho}{\kappa} > 1 \Rightarrow \ln \frac{\rho}{\kappa} > 0$$

$$\frac{\rho}{\kappa} = 1 \Rightarrow \ln \frac{\rho}{\kappa} = 0$$

$$\frac{\rho}{\kappa} < 1 \Rightarrow \ln \frac{\rho}{\kappa} < 0.$$

110 Interestingly, the absolute values of these two parameters are not important, only their ratio. The reaction
 111 coefficient ρ determines the “speed” of the reaction, so that the larger its value, the more efficient is the reac-
 112 tion. The diffusion coefficient κ can be compared to the diffusion parameter of a particule or a molecule as
 113 considered in the early developments of modern physics (Einstein, 1905; Sutherland, 1905) and determines
 114 how fast the molecules move in the environment. Therefore, it makes sense that the ratio ρ/κ is important
 115 as it quantifies the opposite actions of reaction and diffusion. Third and finally, temperature has an almost
 116 linear effect on \mathfrak{G} (Fig. 2). The ratio ρ/κ influences the position of this quasilinear relationship: the higher
 117 this ratio, the lower the value of \mathfrak{G} for a given value of T . The predicted values for \mathfrak{G} vary between 65 and
 118 90 kJ.mol⁻¹ for temperatures varying between 270 and 330 K and a ratio ρ/κ between 0.1 and 10.

119 2.2. Self-Replicating RNA and Information

120 In this section, we consider a real system that might conform to the previous theoretical model. Orgel
 121 (1992) discussed the case of self-replicating RNA molecules with the particular case of the hexamer GC-
 122 CGGC. This single-stranded RNA molecule can self-replicate *in vitro* through the ligation of the two trin-
 123 ucleotides GCC and GGC (Sievers and von Kiedrowski, 1994). Interestingly, this hexameric molecule is
 124 palindromic: it is identical to its Watson–Crick complement so that a pool of these molecules can form
 125 double stranded RNAs, a fact which plays a role in its self-replication process (Orgel, 1992). Another in-
 126 teresting fact about palindromic nucleic acid sequences is that they may form Watson–Crick pairs within
 127 themselves (i.e., hairpins) if they are sufficiently long. The sequence GCCGGC is too short to form a hair-
 128 pin (Fig. 3A); however, the insertion of two nucleotides, for instance GC, in the middle of this hexamer,
 129 is enough to result in a molecule which is predicted to make a hairpin (Fig. 3B). The resulting hairpin is
 130 made of two base pairs. Inserting two other nucleotides C and G result in a decamer which also forms
 131 a hairpin this time with three base pairs (Fig. 3C). These two hypothetical octomer and decamer are also
 132 palindromic. Nevertheless, whether these longer RNA molecules are able to self-replicate in the same way
 133 than the GCCGGC sequence is unknown.

134 The hairpin structure of single-stranded RNAs is important because it helps to stabilize these molecules
 135 by lowering their energy level. Furthermore, these structures are reminiscent of those observed in ribozymes

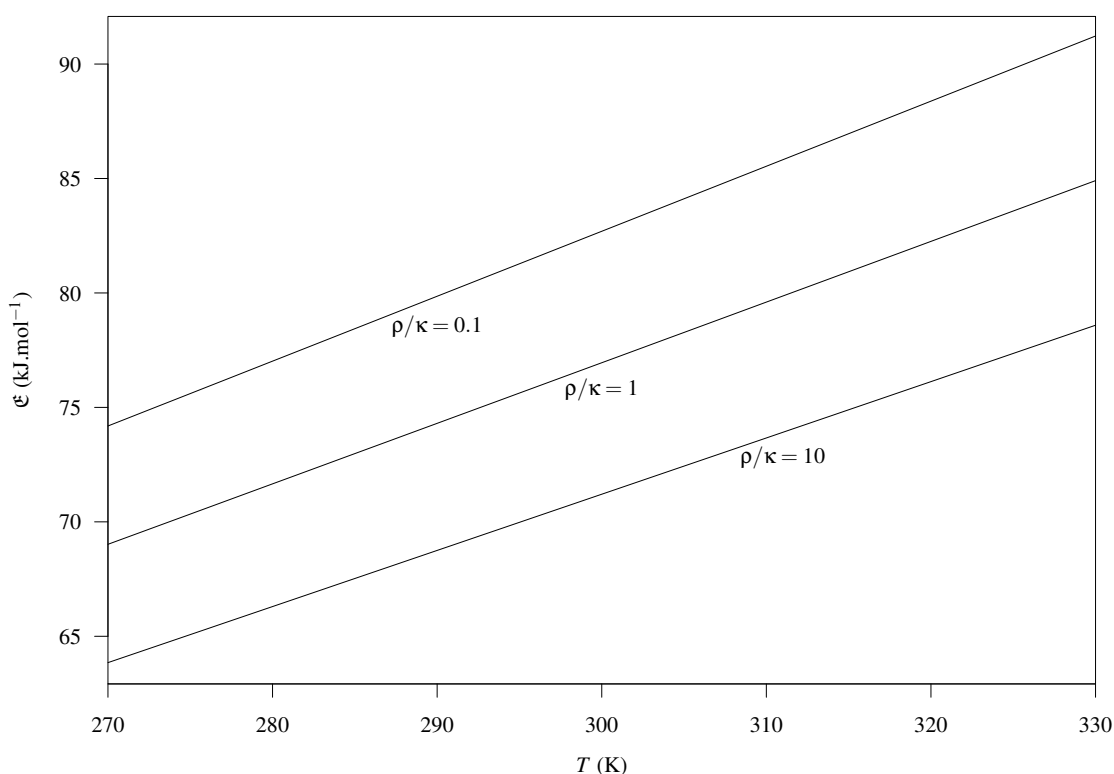


Figure 2: Variation of the equivalent energy of replication (\mathcal{E} in kilojoules per mole) with respect to temperature (T in kelvins) for three different values of the ratio of reaction on diffusion coefficients (ρ/κ). The values were calculated with $Z = 2$ and the values of the constants as given in the text.

136 (Kruger et al., 1982). Indeed, in ribozymes the hairpins and loops making the secondary and tertiary struc-
 137 tures are crucial for their enzymatic activities (Lilley, 2011). Furthermore, Lincoln and Joyce (2009) showed
 138 that ribozymes can be self-replicating.

139 2.3. Information and the evolution of complexity

140 The previous section considers the information–energy equivalence from the standpoint of replication.
 141 The present section starts from this relation between information and energy and considers some of the
 142 consequences in terms of information contents. A hypothetical living system with information stored on a
 143 single bit (like the one considered in the previous section) cannot code more than two states. However, if
 144 the information storage is made of several bits, it could be that more complex information can be coded.
 145 Take the simple case of a sequence with two bits; there are two ways to consider the information coded by
 146 this sequence. The first way is to assume that each bit codes separately for one information each, thus, with
 147 a single bit with a energy-equivalent of $2k_B T \ln 2$.

148 The second way to consider the information coded by two bits is that they may also interact so that the
 149 four possible states (00, 01, 10, 11) may also content new information. The energy of these two interacting
 150 bits can be calculated similarly as they code for a four-state information content by substituting $\ln 2$ by $\ln 4$:

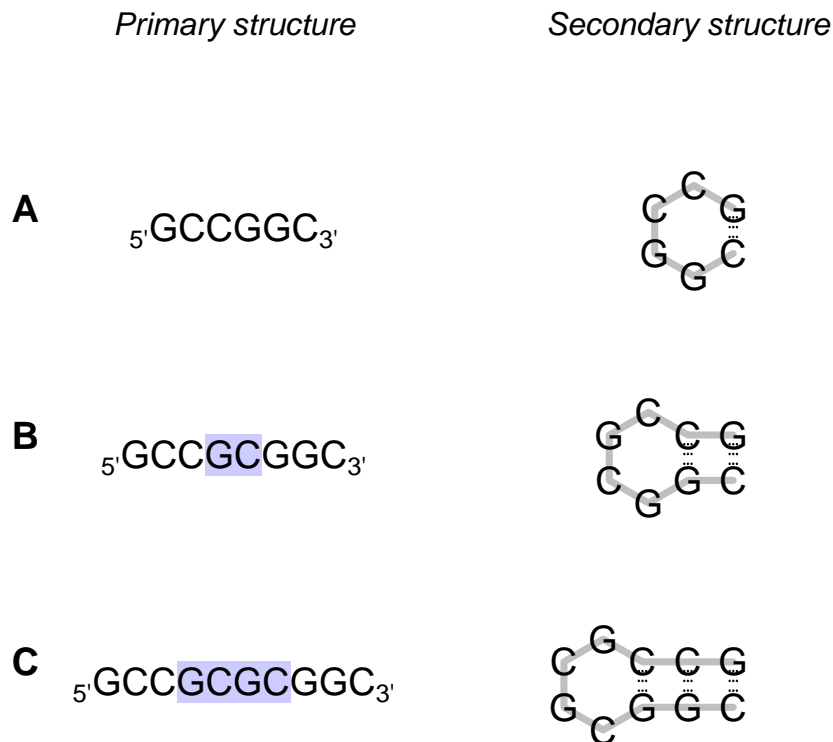


Figure 3: (A) This sequence of six nucleotides can self-replicate (see Orgel, 1992) and is predicted to make a single loop. The insertion of (B) two or (C) four nucleotides results in a more complex secondary structure with a loop and a hairpin. The inserted nucleotides are highlighted on a light blue background. The secondary structures were predicted from the primary structures using <http://rtools.cbrc.jp/>.

$$k_B T \ln 4 = 2k_B T \ln 2.$$

151 It might seem logical that the two ways just described are exclusive. Two bits of information may be
 152 either additive or interactive, but cannot be both. However, in a general approach of modelling interactions
 153 and complexity in the context of information storage, we consider that this last possibility (“both”) is not
 154 excluded. This approach makes possible to define an upper bound to the information contents of a sequence
 155 of bits. Therefore, the overall energy–information content of a sequence of two bits is given by:

$$\Delta E \geq 4k_B T \ln 2.$$

156 We now consider the more general case of a sequence made of n bits. Each bit has a binary information

157 content so that the total information content of this sequence is $nk_{\text{B}}T \ln 2$. Let us assume that some of
 158 these bits interact to create new information. We introduce a new parameter, denoted as ξ , defined as the
 159 proportion of bits that interact to produce new information content ($0 \leq \xi \leq 1$). These interactions could
 160 be of order two (i.e., involving two bits) or of higher orders. We will thus write this new parameter as a
 161 vector ξ_o where the subscript gives the order of the interactions ($o = 2, \dots, N$). Let us define the potential
 162 sequence information (denoted as Ψ) by:

$$\begin{aligned} \Psi &= nk_{\text{B}}T \ln 2 + \sum_{o=2}^N \frac{\xi_o n}{o} k_{\text{B}}T \ln 2^o \\ &= nk_{\text{B}}T \ln 2 \left(1 + \sum_{o=2}^N \xi_o \right). \end{aligned}$$

163 Thus, the additional information content provided by interactions among the bits in a sequence can be
 164 increased by the proportion of interacting bits.

165 We consider now a specific model of two bits with a special type of interaction: regulation. We assume
 166 that the first bit codes for a phenotype (hereafter denoted as either ϕ_0 or ϕ_1), whereas the second bit codes
 167 for the regulation of the expression of this phenotype. Using the same notation as above, we write:

$$x = \begin{cases} 0 \rightarrow \phi_0 \\ 1 \rightarrow \phi_1 \end{cases}$$

168 The second bit is denoted as r and regulates whether x is expressed or not:

$$r = \begin{cases} 0 \rightarrow x \text{ is repressed} \\ 1 \rightarrow x \text{ is expressed} \end{cases}$$

169 We have now a phenotype space for each possible sequence of these two bits:

$$xr = \begin{cases} 00 \rightarrow \emptyset \\ 10 \rightarrow \emptyset \\ 01 \rightarrow \phi_0 \\ 11 \rightarrow \phi_1 \end{cases}$$

170 However, to be effective the regulatory system needs to acquire information, for instance from the environ-
 171 ment. If the information acquired is binary, the associated energy cost is at least $k_B T \ln 2$. Consequently,
 172 the energy cost of the phenotypic expression requires to acquire three bits of information: one bit from the
 173 environment and the two bits from the sequence. The energy equivalent is:

$$\Delta E \geq 3k_B T \ln 2.$$

174 In the general case, if there are n_x bits for the phenotypes, and n_r bits for the regulatory systems:

$$\begin{aligned} \Delta E &\geq (n_x + n_r)k_B T \ln 2 + n_r k_B T \ln 2 \\ &\geq (n_x + 2n_r)k_B T \ln 2. \end{aligned}$$

175 So the energetic costs of regulatory systems is predicted to be twice larger than for phenotype coding
 176 sequences.

177 3. Discussion

178 Information theory is concerned with the mechanisms of information transmission and interpretation
 179 (e.g., coding and decoding; Cover and Thomas, 2006). Among the many applications of information theory
 180 in physical sciences, the Landauer's principle predicts that information acquisition and erasure in comput-
 181 ing machines entail a rise in entropy, therefore linking thermodynamics and information theory (Landauer,
 182 1961). Experimental evidence has been recently found to support this conjecture (Toyabe et al., 2010; Bérut
 183 et al., 2012; Hong et al., 2016), although the interpretation of these experimental results has been disputed
 184 (Kish and Granqvist, 2012). As an extension of Landauer's principle, theoretical physicists recently at-
 185 tempted to quantify the mass of a bit of information (Herrera, 2014; Vopson, 2019), therefore effectively
 186 linking information theory with special relativity. Although the idea of entropy seems ubiquitous enough
 187 to be used in different contexts, it is important to differentiate clearly those applications. Recently, Vopson
 188 and Lepadatu (2022) clarified this point by stating the law of infodynamics where the total entropy of a
 189 system is defined as the sum of its physical entropy (in its traditional thermodynamic meaning as the sta-
 190 tistical distribution of microstates) and its information entropy (defined with Shannon's classic formula).
 191 Their law has been applied to analyse genetic mutations of the SARS-CoV2 viruses within a deterministic
 192 framework (Vopson, 2022). When considering biological systems, the distinction between both forms of
 193 entropy is indeed important. Thermodynamic entropy considers the spatial distribution of entities which can

194 differ in any relevant characteristic (e.g., the velocity of gas molecules in Maxwell's hypothetical demon
195 experiment) but this quantity has no implicit or explicit consequence in terms of information. On the other
196 hand, Shannon's entropy considers the relative frequencies of these entities (i.e., their diversity) as potential
197 information bearers. Both concepts are crucial for biological processes and cannot be used interchangeably.

198 The present paper starts from Brillouin's equation of information to assess the energy requirements of
199 two fundamental aspects of living systems: replication and phenotype expression. The goal of this work
200 is to find the conditions or requirements that permitted the emergence of life even though the underlying
201 physical and chemical processes follow the second law of thermodynamics.

202 The problem of replication of living beings, and in particular the replication of genomic sequences, is
203 the perfect illustration that the chemical reactions involved in this process do not respect the second law of
204 thermodynamics, by definition, since genomes are replicated identical to themselves resulting in a decrease
205 in entropy of the system. The challenging question is: How this process was initiated during evolution?
206 The model developed in this paper proposes a mechanism based on the energy requirements involved in
207 the reactions of replication: the process of information recognition during replication lowers the internal
208 energy of the system and can provide the necessary energy for replication and information duplication.

209 Initially during evolution, there was no blueprint similar to the current genomic information, so repli-
210 cation could not occur by producing identical individuals. Instead, replication could have happened by
211 mechanical or physical replication. Such mechanisms do not involve any biological processes; however,
212 they are known to favour the replication of some entities by simple preferential pairing of elements (e.g.,
213 Penrose and Penrose, 1957; Griffith et al., 2005). This mechanical process of self-replication can be re-
214 lated to some results on the replication of small RNA sequences (Orgel, 1992; Sievers and von Kiedrowski,
215 1994).

216 It is interesting to note that the synthesis of small RNA molecules can be achieved through coupled
217 cyclical reactions (Sievers and von Kiedrowski, 1994; Robertson and Joyce, 2012). Liu and Sumpter (2018)
218 reviewed coupled chemical reactions and suggested that they may have had an important role in the early
219 developments of life. One prediction from the present paper is that short, self-replicating RNA sequences
220 may have been synthesized by accumulating nucleotides and result in molecules with enzymatic activities
221 similar to those observed in current ribozymes (Kruger et al., 1982; Lincoln and Joyce, 2009; Lilley, 2011).
222 It has been shown that even some very short RNA sequences, up to five nucleotides, can have an enzymatic
223 activity (Turk et al., 2010). It remains to be shown that the short sequences that are able to self-replicate
224 are also able to have enzymatic activities. That would bridge the gap between replication and (bio)chemical
225 activity.

226 In living beings, information stored in genomes can be interpreted as a device providing energy since,
227 in the case of enzyme-coding sequences, the protein that are coded and eventually synthesized contribute to

228 lower the activation energy of the reactions in the cell (Ji, 1974, 2000). Therefore, there is a conversion from
229 information into energy. Before the evolution of genomes, this flow could not have happened in this direction,
230 and it seems we need to hypothesize that energy was converted into information in a way or another.

231 Since the “RNA World” hypothesis has been formulated, several studies have provided a range of evi-
232 dence for its support (Robertson and Joyce, 2012). Similarities between the sequences of different types of
233 RNA found in the cells have been noted by some authors several decades ago. Even before the discovery
234 of the ribozyme, Mizutani and Ponnampereuma (1980) noted “that the ancestral form of tRNA’s could have
235 been composed of 17 nucleotides of which 7 nucleotides formed a loop and of which two pentanucleotides
236 at either end of the loop constituted complementary double helical structure.” However, these authors did
237 not favour the hypothesis of “RNA first” and thought that “it is reasonable to suppose that small peptides
238 and small oligonucleotides as well as their precursors existed and interacted in the ‘soup’.” See also White
239 (1980) who hypothesized the peptides and nucleotide oligomers interacted closely to generate the first living
240 cells.

241 The discovery of enzymatic activity in some RNA molecules changed the perspective that initial living
242 beings had to rely on proteins. Lacey et al. (1992) reviewed some of the literature following this discovery
243 and pointed out to the similarity between the sequences and structures of ribozymes and tRNAs: “Maybe 5’-
244 CA-3’ is one of the simpler (and most important) ribozymes and maybe AMP by itself is just not sufficient.
245 We shall see.” These authors also noted that similarities between the different types of RNA pointed to a
246 common origin of these molecules: “More recent homology studies (Staves and Lacey, 1989) support the
247 idea of a common ancestor for tRNAs and 5S rRNA.” A recent review by Tamura (2015) seems to confirm
248 that rRNA could have evolved from ancestral tRNA. Finally, Becker et al. (2019) showed that all purines,
249 pyrimidines, and ribose (i.e., the building blocks of ribonucleotides and then of RNA) can be synthesized
250 in realistic prebiotic conditions driven by wet-dry cycles, further supporting that RNA-based living entities
251 may have been the first to evolve on Earth.

252 These findings could seem intriguing given the fact that current rRNA and tRNA molecules interact
253 closely with many proteins. For instance, the ribosome is made of several rRNA molecules and several
254 proteins. However, it has been shown that a ribosome made of rRNAs only can be functional:

255 The most obvious source of energy comes from peptide bond formation; peptidyl transfer re-
256 sults in formation of a peptide amide bond from an activated ribose ester linkage, accompanied
257 by a large change in free energy. How can this free energy change be coupled to transloca-
258 tion? The answer most likely lies in the changing chemical nature of the acceptor end of the
259 tRNA as it moves through the ribosome (Spirin 1985). [...] This finding shows that ther-
260 mal energy alone is sufficient to drive the intersubunit rotation underlying translocation. [...]

261 These findings show that even the complex, large-scale molecular movements associated with
262 translocation can be driven by thermal energy, obviating the need for special energy-generating
263 steps in protein synthesis by the first ribosomes. (Noller, 2012)

264 It appears therefore that evolution may have started well before DNA genomes appeared. In a more
265 recent study, Goldman et al. (2010) found that even RNA genomes may have interacted and coevolved with
266 proteins that are still in use in cells today. In particular, they found that “translation proteins underwent
267 major evolutionary expansion well before the first species diverged from LUCA (last universal common
268 ancestor) and even before the DNA genome was established.”

269 The present work agrees with the RNA World hypothesis since it predicts that an entity can be replicated
270 where the required energy is provided by the lower energy level of the system during information reading.
271 Self-replicating RNA molecules seem a good empirical model of such entities.

272 Another line of evidence is found from coupled chemical reactions. In living cells, coupled reactions
273 occur frequently so that a thermodynamically favoured reaction (e.g., hydrolysis of ATP into ADP) releases
274 energy for a second reaction which is thermodynamically less favoured (e.g., binding of an amino acid to
275 a peptide). Liu and Sumpter (2018) reviewed models of coupled chemical reactions and found that some
276 systems are catalytic while others are self-replicating. More importantly, they found that “it is possible to
277 construct a self-replicating system in which the entropy of some parts spontaneously decreases.” Recent
278 works have proposed several experimental systems of self-replicating chemical systems that could be mod-
279 els of prebiotic evolution (Komáromy et al., 2017; Colomer et al., 2018; Morrow et al., 2019; Liu et al.,
280 2020). It is still not yet clear whether one of these systems could be the precursor of biological evolution
281 on Earth, or whether several of these systems could have coexisted on Earth before one of them (based on
282 RNA self-replicating molecules) evolved.

283 Darwinian evolution is dependent on the existence of information storage in living beings (Dawkins,
284 1996). Therefore, natural selection certainly did not play a role in the early stages of life evolution when
285 information was not firmly established. However, once self-replicating entities with information storage
286 have evolved, selection among them could have acted and favoured the replication of some of them. I draw
287 here a similarity between the mechanisms of regulation and Maxwell’s demon (Earman and Norton, 1998;
288 Weinstein, 2003). As a reminder, a demon is, according to Maxwell, an entity which is able to decrease the
289 entropy of a system without inputting energy in it by opening a gate to let the “fast” (i.e., high energy) gaz
290 particules move in one direction and the “slow” (low energy) ones move in the other direction. However,
291 Szilard (1964) showed that the demon must acquire information about their energy levels, and given the
292 equivalence stated in the Introduction, this requires to spend some energy. Furthermore, in Maxwell’s
293 riddle, the demon is given. We could see an analogy with living systems where regulatory mechanisms act

294 effectively as demons by decreasing the entropy. However, Maxwell's original riddle does not say anything
295 how the demon acquired the information that tells it to do its work. Indeed, the demon must have been
296 "programmed" in one way or another. Similarly, in living systems the emergence of regulation during the
297 early stages of biological evolution faces the same problem. How come these mechanisms have come to do
298 what they do—regulate biological processes?

299 An argument of the present paper is that considering the equivalence between energy and information
300 makes possible to explain the origin of complexity from the same principle that explains the emergence
301 of self-replicating systems with information storage. Zuckerkandl (2006) emphasized the importance of
302 higher-order interactions in biological processes, stating that information stored in DNA sequences is, in
303 itself, not sufficient:

304 The requirement for an interaction between sequence properties and higher-order structural
305 properties of informational macromolecules however may well remain absolute for all forms
306 of life in the universe. (p. 6)

307 Zuckerkandl also noted that a "linear heteropolymer acting alternatively through its sequence and through its
308 conformation-dependent enzymatic properties" can fulfill the dual role of information storage and biological
309 activity.

310 The energy requirements for biochemical reactions during the first stages of evolution have been an
311 unanswered question for several decades. Because of this difficulty, several authors formulated the hypoth-
312 esis of panspermia which postulates that life may have been originated outside the Earth and brought on our
313 planet through meteorites or other extraterrestrial vehicles (Arrhenius, 1908; Crick and Orgel, 1973). How-
314 ever, Orgel (1973) proposed that energy was brought to early biochemical reactions through an alternance
315 of dry and wet environments:

316 If dry thermal reactions were important for the origins of life, it is hard to understand how they
317 were replaced by solution reactions later in the course of biochemical evolution.

318 I believe that the two types of reaction occurred in close association on the primitive earth.
319 Random polymers and reactive intermediates were formed in efficient thermal reactions. When
320 the mixtures of products formed in this way were wetted, for example by rain or dew, they
321 reacted further in solution. If this picture is correct, modern biological condensations evolved
322 from these latter solution reactions. (p. 138)

323 Some minerals are known to catalyze RNA polymerization (for instance, montmorillonites; Aldersley
324 et al., 2011) and could have played a role in the early self-replication of short RNA molecules (Becker et al.,
325 2019). These catalytic minerals may have been important for the early synthesis of RNA polymers before

326 they could be used as information storage devices. Once these polymers were self-replicating, so that energy
327 could be stored as information, the early stages of evolution could be initiated. These speculations can be
328 related to a prediction from the above model which is that the higher the ratio of the reaction coefficient to
329 the diffusion coefficient (ρ/κ), the lower the energy requirement for replication. Earlier researchers have
330 noted the importance of “keeping together” the molecules involved in the first stages of prebiotic evolution:

331 It is important to note that autocatalytic growth and natural selection cannot occur in this model
332 unless both peptide and nucleotide oligomers remain in the vicinity of each other much of the
333 time, either due to the impermeability of a membrane, to adsorption on a clay surface, or to
334 some similar localization process. (White, 1980, p. 126)

335 By contrast to White’s model which postulates that proteins and nucleic acids were both required during
336 prebiotic evolution, the hypothesis of RNA molecules associated with, for instance, clay minerals provides
337 a simpler mechanism because it does not require the early presence of proteins.

338 The model developed in this paper leads to the prediction that the energy requirement for replication
339 of a bit varies between 65 and 90 $\text{kJ}\cdot\text{mol}^{-1}$ for temperatures varying between 270 and 330 K (0–60 °C).
340 Such values are close to the decrease in energy resulting from the formation of H-bonds in DNA sequences
341 which has been estimated to be 52 and 106 $\text{kJ}\cdot\text{mol}^{-1}$ for AT and GC pairs, respectively (Mo, 2006) which
342 support the argument developed above. Additionally, it is predicted that binary sequences, i.e. with the
343 parameter $Z = 2$, have a lower energy requirement than with $Z > 2$. This suggests the possibility that
344 initial self-replicating RNA sequences might have evolved from GC pairs only, and that AU pairs may have
345 been incorporated later during evolution. The small hypothetical RNA molecules considered above with
346 eight or ten nucleotides are indeed made only of G and C and are predicted to form a loop and a hairpin.
347 It will be interesting to assess whether some RNA molecules made only of these two bases have catalytic
348 activity.

349 By providing a theoretical framework, the above model can be applied to the origin of life in other
350 environments than on Earth. The above equations were derived without assumption on the exact nature of
351 the molecules involved. Thus, any system that agrees with the above predictions might lead to the evolution
352 of life. It will be particularly interesting to assess whether this model can have applications for the study of
353 extraterrestrial life.

354 **Declaration of competing interest**

355 The author declares that he has no known competing financial interests or personal relationships that
356 could have appeared to influence the work reported in this paper.

357 **Acknowledgements**

358 I am grateful to the anonymous reviewer for constructive comments. This is publication ISEM 2023-
359 051.

360 **References**

- 361 Aldersley, M.F., Joshi, P.C., Price, J.D., Ferris, J.P., 2011. The role of montmorillonite in its catalysis of RNA synthesis. *Appl. Clay*
362 *Sci.* 54, 1–14. doi:10.1016/j.clay.2011.06.011.
- 363 Arrhenius, S., 1908. *Worlds in the making. The evolution of the Universe.* Harper & Brothers, New York and London.
- 364 Becker, S., Feldmann, J., Wiedemann, S., Okamura, H., Schneider, C., Iwan, K., Crisp, A., Rossa, M., Amatov, T., Carell, T., 2019.
365 Unified prebiotically plausible synthesis of pyrimidine and purine RNA ribonucleotides. *Science* 366, 76–82. doi:10.1126/
366 science.aax2747.
- 367 Bekenstein, J.D., 1973. Black holes and entropy. *Phys. Rev. D* 7, 2333–2346. doi:10.1103/physrevd.7.2333.
- 368 Bérut, A., Arakelyan, A., Petrosyan, A., Ciliberto, S., Dillenschneider, R., Lutz, E., 2012. Experimental verification of Landauer's
369 principle linking information and thermodynamics. *Nature* 483, 187–189. doi:10.1038/nature10872.
- 370 Bonnici, V., Manca, V., 2016. Informational laws of genome structure. *Sci. Rep.* 6, 28840. doi:10.1038/srep28840.
- 371 Brillouin, L., 1949. Life, thermodynamics, and cybernetics. *Am. Sci.* 37, 554–568.
- 372 Brillouin, L., 1953. The negentropy principle of information. *J. Appl. Phys.* 24, 1152–1163. doi:10.1063/1.1721463.
- 373 Brillouin, L., 1959. Information theory and its applications to fundamental problems in physics. *Nature* 183, 501–502. doi:10.
374 1038/183501a0.
- 375 Colomer, I., Morrow, S.M., Fletcher, S.P., 2018. A transient self-assembling self-replicator. *Nature Comm.* 9. doi:10.1038/
376 s41467-018-04670-2.
- 377 Cover, T.M., Thomas, J.A., 2006. *Elements of information theory (second edition).* John Wiley & Sons, Hoboken, NJ.
- 378 Crick, F., 1981. *Life itself. Its origin and nature.* Simon and Schuster, New York.
- 379 Crick, F.H.C., Orgel, L.E., 1973. Directed panspermia. *Icarus* 19, 341–346. doi:https://doi.org/10.1016/
380 0019-1035(73)90110-3.
- 381 Dawkins, R., 1996. *The blind watchmaker.* W. W. Norton & Company, New York.
- 382 Earman, J., Norton, J.D., 1998. Exorcist XIV: The wrath of Maxwell's Demon. Part I. From Maxwell to Szilard. *Stud. Hist. Philos.*
383 *Modern Phys.* 29, 435–471. doi:10.1016/S1355-2198(98)00023-9.
- 384 Einstein, A., 1905. Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten
385 suspendierten Teilchen. *Ann. Phys.* 322, 549–560. doi:10.1002/andp.19053220806.
- 386 Elliott, T.A., Gregory, T.R., 2015. What's in a genome? The C-value enigma and the evolution of eukaryotic genome content. *Phil.*
387 *Trans. R. Soc. Lond. B* 370, 20140331. doi:10.1098/rstb.2014.0331.
- 388 Eyring, H., 1935. The activated complex in chemical reactions. *J. Chem. Phys.* 3, 107–115. doi:10.1063/1.1749604.
- 389 Gibbs, J.W., 1873. A method of geometrical representation of the thermodynamic properties of substances by means of surfaces.
390 *Transactions of the Connecticut Academy of Arts and Sciences* 2, 382–404. URL: https://www3.nd.edu/~powers/ame.
391 20231/gibbs1873b.pdf.
- 392 Goldman, A.D., Samudrala, R., Baross, J.A., 2010. The evolution and functional repertoire of translation proteins following the origin
393 of life. *Biol. Direct* 5, 15. doi:10.1186/1745-6150-5-15.
- 394 Griffith, S., Goldwater, D., Jacobson, J.M., 2005. Self-replication from random parts. *Nature* 437, 636–636. doi:10.1038/

- 395 437636a.
- 396 Hawking, S.W., Ellis, G.F.R., 1973. The large scale structure of space–time. Cambridge University Press, Cambridge.
- 397 Herrera, L., 2014. The mass of a bit of information and the Brillouin’s principle. *Fluct. Noise Lett.* 13, 1450002. doi:10.1142/
398 S0219477514500023.
- 399 Hong, J., Lambson, B., Dhuey, S., Bokor, J., 2016. Experimental test of Landauer’s principle in single-bit operations on nanomagnetic
400 memory bits. *Sci. Adv.* 2, e1501492. doi:10.1126/sciadv.1501492.
- 401 Ji, S., 1974. Energy and negentropy in enzymic catalysis. *Ann. N.Y. Acad. Sci.* 227, 419–437. doi:10.1111/j.1749-6632.
402 1974.tb14405.x.
- 403 Ji, S., 2000. Free energy and information contents of Conformons in proteins and DNA. *BioSystems* 54, 107–130. doi:10.1016/
404 S0303-2647(99)00073-8.
- 405 Kish, L.B., Ferry, D.K., 2018. Information entropy and thermal entropy: apples and oranges. *J. Comput. Electron.* 17, 43–50.
406 doi:10.1007/s10825-017-1044-1.
- 407 Kish, L.B., Granqvist, C.G., 2012. Energy requirement of control: comments on Szilard’s engine and Maxwell’s demon. *Europhys.*
408 *Lett.* 98, 68001. doi:10.1209/0295-5075/98/68001.
- 409 Komáromy, D., Stuart, M.C.A., Santiago, G.M., Tezcan, M., Krasnikov, V.V., Otto, S., 2017. Self-assembly can direct dynamic
410 covalent bond formation toward diversity or specificity. *J. Am. Chem. Soc.* 139, 6234–6241. doi:10.1021/jacs.7b01814.
- 411 Kruger, K., Grabowski, P.J., Zaug, A.J., Sands, J., Gottschling, D.E., Cech, T.R., 1982. Self-splicing RNA: autoexcision and auto-
412 cyclization of the ribosomal RNA intervening sequence of Tetrahymena. *Cell* 31, 147–157. doi:10.1016/0092-8674(82)
413 90414-7.
- 414 Lacey, J.C., Wickramasinghe, N.S.M.D., Cook, G.W., 1992. Experimental studies on the origin of the genetic code and the process of
415 protein synthesis: a review update. *Origins Life Evol. Biosph.* 22, 243–275. doi:10.1007/bf01810856.
- 416 Landauer, R., 1961. Irreversibility and heat generation in the computing process. *IBM Journal* 5, 183–191. doi:10.1147/rd.53.
417 0183.
- 418 Lilley, D.M.J., 2011. Mechanisms of RNA catalysis. *Phil. Trans. R. Soc. Lond. B* 366, 2910–2917. doi:10.1098/rstb.2011.
419 0132.
- 420 Lincoln, T.A., Joyce, G.F., 2009. Self-sustained replication of an RNA enzyme. *Science* 323, 1229–1232. doi:10.1126/science.
421 1167856.
- 422 Liu, B., Pappas, C.G., Ottel , J., Schaeffer, G., Jurissek, C., Pieters, P.F., Altay, M., Mari , I., Stuart, M.C.A., Otto, S., 2020.
423 Spontaneous emergence of self-replicating molecules containing nucleobases and amino acids. *J. Am. Chem. Soc.* 142, 4184–
424 4192. doi:10.1021/jacs.9b10796.
- 425 Liu, Y., Sumpter, D.J.T., 2018. Mathematical modeling reveals spontaneous emergence of self-replication in chemical reaction sys-
426 tems. *J. Biol. Chem.* 293, 18854–18863. doi:10.1074/jbc.ra118.003795.
- 427 Lynch, M., Conery, J.S., 2003. The origins of genome complexity [Letters: vol. 304, pp. 389–390]. *Science* 302, 1401–1404.
- 428 Manca, V., 2017. The principles of informational genomics. *Theor. Comput. Sci.* 701, 190–202. doi:10.1016/j.tcs.2017.02.
429 035.
- 430 Mizutani, H., Ponnampereuma, C., 1980. The evolution of the protein synthesis system, II. From chemical evolution to biological
431 evolution. *Origins of Life* 10, 31–38. doi:10.1007/bf00928941.
- 432 Mo, Y., 2006. Probing the nature of hydrogen bonds in DNA base pairs. *J. Mol. Model.* 12, 665–672. doi:10.1007/
433 s00894-005-0021-y.
- 434 Morrow, S.M., Colomer, I., Fletcher, S.P., 2019. A chemically fuelled self-replicator. *Nature Comm.* 10. doi:10.1038/

- 435 s41467-019-08885-9.
- 436 Noller, H.F., 2012. Evolution of protein synthesis from an RNA world. *Cold Spring Harbor Persp. Biol.* 4, a003681. doi:10.1101/
- 437 cshperspect.a003681.
- 438 Orgel, L.E., 1973. *The origins of life. Molecules and natural selection.* John Wiley & Sons, Hoboken, NJ.
- 439 Orgel, L.E., 1992. Molecular replication. *Nature* 358, 203–209. doi:10.1038/358203a0.
- 440 Paradis, E., 2020. *Population genomics with R.* Chapman & Hall, Boca Raton, FL.
- 441 Penrose, L.S., Penrose, R., 1957. A self-reproducing analogue. *Nature* 179, 1183–1183. doi:10.1038/1791183a0.
- 442 Robertson, M.P., Joyce, G.F., 2012. The origins of the RNA world. *Cold Spring Harbor Persp. Biol.* 4, a003608. doi:10.1101/
- 443 cshperspect.a003608.
- 444 Shannon, C.E., 1948. A mathematical theory of communication. *Bell System Tech. J.* 27, 379–423, 623–656. doi:10.1002/j.
- 445 1538-7305.1948.tb01338.x.
- 446 Sievers, D., von Kiedrowski, G., 1994. Self-replication of complementary nucleotide-based oligomers. *Nature* 369, 221–224. doi:10.
- 447 1038/369221a0.
- 448 Sutherland, W., 1905. A dynamical theory of diffusion for non-electrolytes and the molecular mass of albumin. *London, Edinburgh,*
- 449 *and Dublin Philosophical Magazine and Journal of Science* 9, 781–785. doi:10.1080/14786440509463331.
- 450 Szilard, L., 1964. On the decrease of entropy in a thermodynamic system by the intervention of intelligent beings. *Behav. Sci.* 9,
- 451 301–310. doi:10.1002/bs.3830090402.
- 452 Tamura, K., 2015. Origins life evol. biosph. *Life* 5, 1687–1699. doi:10.3390/life5041687.
- 453 Toyabe, S., Sagawa, T., Ueda, M., Muneyuki, E., Sano, M., 2010. Experimental demonstration of information-to-energy conversion
- 454 and validation of the generalized Jarzynski equality. *Nature Phys.* 6, 988–992. doi:10.1038/NPHYS1821.
- 455 Truhlar, D.G., Hase, W.L., Hynes, J.T., 1983. Current status of transition-state theory. *J. Chem. Phys.* 87, 2664–2682. doi:10.1021/
- 456 j100238a003.
- 457 Turk, R.M., Chumachenko, N.V., Yarus, M., 2010. Multiple translational products from a five-nucleotide ribozyme. *Proc. Natl. Acad.*
- 458 *Sci. USA* 107, 4585–4589. doi:10.1073/pnas.0912895107.
- 459 Vitas, M., Dobovišek, A., 2019. Towards a general definition of life. *Origins Life Evol. Biosph.* 49, 77–88. doi:10.1007/
- 460 s11084-019-09578-5.
- 461 Vopson, M.M., 2019. The *mass-energy-information* equivalence principle. *AIP Adv.* 9, 095206. doi:10.1063/1.5123794.
- 462 Vopson, M.M., 2022. A possible information entropic law of genetic mutations. *Appl. Sci.* 12, 6912. doi:10.3390/app12146912.
- 463 Vopson, M.M., Lepadatu, S., 2022. Second law of information dynamics. *AIP Adv.* 12, 075310. doi:10.1063/5.0100358.
- 464 Weinstein, S., 2003. Objectivity, information, and Maxwell’s demon. *Philos. Sci.* 70, 1245–1255. doi:10.1086/377404.
- 465 White, D.H., 1980. A theory for the origin of a self-replicating chemical system. I: Natural selection of the autogen from short, random
- 466 oligomers. *J. Mol. Evol.* 16, 121–147. doi:10.1007/bf01731582.
- 467 Zuckerkandl, E., 2006. Intelligent design and biological complexity. *Gene* 385, 2–18. doi:10.1016/j.gene.2006.03.025.