Journal of Physical Chemistry & Biophysics
ISSN: 2161-0398

Otsuka, J Phys Chem Biophys 2017, 7:1
DOI: 10.4172/2161-0398.1000235

Research Article OMICS International

The Concept of Biological Activity and Its Application to Biological Phenomena

Otsuka J*

JO Institute of Biophysics, Higashi-Ohizumi, Nerima-ku, Tokyo, Japan

Abstract

Although the life has been a mystery for most physicists since the problem of Maxwell's demon, this mystery is resolved by considering the following characteristics of an organism; the self-reproduction by taking material and energy sources from the outside on the basis of its genetic information, and the selection of self-reproduced organisms to maintain and further improve the genetic information. According to the knowledge of molecular biology revealed recently, the molecular route to accomplish the self-reproduction is evaluated energetically, and a new thermodynamic quantity of biological activity is proposed for characterizing the state of an organism in terms of acquired energy, stored energy and systematization. This quantity is not only compatible with the law of thermodynamics but also reflects the changes in genome and in the mode of gene expression. Thus, the biological activity becomes a useful measure for analyzing various biological phenomena quantitatively. This is illustrated for the large-scale evolution by generating new genes from gene duplication and for the estimation of the energy required for the development of a multicellular organism. The origin of life is also discussed from the aspect of biological activity and the extended view of evolution.

Keywords: Organism; Thermodynamics laws; Self-reproduction; Selection; Entropy; Maxwell's demon

Introduction

How is it possible to understand the life when the whole world is ruled by the second principle of thermodynamics, seemingly pointing towards disorder? This problem is first raised by Maxwell [1], as the question about the presence of a demon sorting the molecules in gaseous phase. Since then, the paradox of Maxwell's demon has been discussed by many physicists. In particular, Szilard [2] suggests the necessity of introducing the concept of 'information' into the analysis of the operation of Maxwell's demon. According to this suggestion, Brillouin [3] has analyzed the operation that at least one quantum of light emitted from an electric torch is scattered by a molecule and is absorbed in the eye of the demon. In this analysis, he shows that the increase in entropy of the demon by accepting the light quantum is greater than the decrease in entropy by acquiring the information about the molecule, concluding that the demon ultimately dies by the repeat of such operation. Thus, this analysis saved the thermodynamic law but did not succeed in resolving the mystery of life. Although the thermodynamics is extended afterwards to treat the irreversible processes far from equilibrium, this approach is directed to evaluate the entropy production arising from the dissipative structure [4] but does not inquire into the problem how an organism maintains and further strengthens its lower entropy structure. This is also the case for the hypercycle proposed by a series of catalytic actions [5]. Any of these approaches to the life by physicists and physico-chemists lacks the consideration of the characteristics of an organism, i.e., taking food (material and energy source), self-reproduction, and the natural selection proposed by Darwin [6]. Although the self-reproduction is discussed in terms of automata containing its program [7], the problem of Maxwell's demon is left untouched.

Recently, the studies of molecular biology have revealed the central dogma in the free-living organism; the proteins are translated from the messenger RNAs (mRNAs) by the aid of ribosomal RNAs (rRNAs) and transfer RNAs (tRNAs), and the three kinds of RNAs are transcribed from the DNA genes, which are replicated upon self-reproduction [8]. These proteins play the central role in acquiring the material and energy sources from the outside and in converting them into the biomolecules for the growth and self-reproduction of the organism. This complexity of the present-day organism raises the debate on the origin of life; "which came first, the protein or the nucleic acid?" The discovery of the self-splicing activity of rRNA intervening sequence [9] strongly suggests a possibility that the life has started from self-replicative RNAs called the RNA replicases [10]. The RNA replicase is hypothetical one, but this is the simplest model to illustrate that the self-reproducing system can behave as Maxwell's demon, retaining its ability through selection [11]. In the present paper, a new thermodynamic quantity of biological activity is proposed to characterize an organism by extending the analysis on the RNA replicase to the central dogma. This quantity itself not only resolves the paradox of Maxwell's demon but also becomes a useful measure for analyzing biological phenomena quantitatively, bridging the gap between physics and biology. The evolution from the RNA replicase to the present-day organisms is also discussed.

Materials and Methods

Proposal of the concept of biological activity

The characters of an organism are generally determined by two internal variables, the size N of its genome (a set of genes) and the systematization S_N of the genome and its products. The systematization is the degree of negative entropy or negentropy, $-S_N$, which should be measured for the special arrangement of nucleotide bases in genes, the special arrangement of amino acid residues in the proteins transmitted from the genes, the metabolic pathways formed by the catalytic

***Corresponding author:** Otsuka J, JO Institute of Biophysics, Higashi-Ohizumi 6-49-18, Nerima-ku, Tokyo, 178-0063, Japan, Tel: 813 3921 1466; E-mail: jin.otsuka@kyj.biglobe.ne.jp

Received January 30, 2017; **Accepted** February 03, 2017; **Published** February 10, 2017

Citation: Otsuka J (2017) The Concept of Biological Activity and Its Application to Biological Phenomena. J Phys Chem Biophys 7: 235. doi: 10.4172/2161-0398.1000235

actions of enzyme proteins, the regulation and control at the levels of translation, transcription and replication of DNAs, the cell structure constructed by the intrinsic property of lipid molecules to form a vesicle of lipid bilayer and by the cell wall to support the lipid vesicle, and for furthering the communication between differentiated cells in the case of a multicellular organism.

The energy acquired by such an organism depends on its genome size N and systematization S_N as well as on the material and energy source M available from the environment. Thus, the acquired energy is expressed to be $E_a(M; N, S_N)$, which is an increasing function of N and S_N as well as M. On the other hand, the organism utilizes the energy and materials to construct the biomolecules for its growth and self-reproduction. The energy $E_s(N, S_N)$ stored in the biomolecules such as polynucleotides, proteins, lipids and cell wall is also another increasing function of N and S_N. These biomolecules to exhibit biological functions have the higher energy than the inorganic compounds, into which they are finally decomposed. The stored energy in these biomolecules should be measured in comparison with the energy of the decomposed state. The difference between the acquired energy and the stored energy, $E_a(M; N, S_N) - E_s(N, S_N)$, is released as heat. If the entropy production by the heat compensates for the entropy reduction, i.e., $-S_N$, due to the systematization, this is consistent with the law of thermodynamics. In the organism, the acquired energy is transiently trapped in ATP and NADPH molecules as chemical energy, and it is gradually consumed in the syntheses of other biomolecules under the guidance of enzyme proteins, without drastic change in temperature T. Thus, the following quantity will be proposed as biological activity BA.

$$BA=E_a(M; N, S_N)-E_s(N, S_N)-TS_N > 0 \quad (1)$$

The positive value of this quantity will be shown in the next section, and it is considered to be proportional to the self-reproducing rate of an organism as the first approximation, because the biological process tends to proceed more smoothly under the larger value of BA. The biological activity has thermodynamic connotation as a departure from equilibrium, but this is in a reverse relation to the free energy which decreases upon any change in a given system by the decrease in internal energy and/or by the increase in entropy. This is due to the property of the genome that is almost constant during one lifetime of an organism but changes enough to increase the biological activity by the evolution during much longer time, as mentioned below.

While RNAs and proteins are steadily transcribed and translated, respectively, to make up for their destruction, the DNA genome also suffers damages, especially in the nucleotide bases. However, such damaged bases are repaired and proofread. By this procedure, nucleotide bases are only substituted with the rate $u \sim 10^{-9}$ per site per year in both prokaryotes and eukaryotes [12,13]. The prokaryote carries the genome, whose size l is within 10^7 base pairs (bp) [14], and it only suffers the base substitution with the probability of $lut=10^{-2}$ after one year. During this period, the prokaryote repeats many times of self-reproduction. Thus, most of the descendants retain the original genome and only a small fraction of descendants receive substituted bases. Although the latter descendants are mostly defective and compelled to extinction, some of them contribute to the evolution as shown later. The multicellular diploid organisms of eukaryotes have expanded their genome size l to 10^8 bp or more [14]. The biological function of each cell in such an organism is seriously influenced only when the nucleotide bases are substituted at the identical sites of homologous chromosomes. This occurs with the probability $l(ut)^2=10^9(10^{-9}\times 10^2)^2=10^{-5}$ during 10^2 years in a cell of a diploid organism such as the human, whose genome consists of two sets of 10^9 bp. Among n cells constituting the adult form of such an organism, therefore, the number of cells not having suffered the seriously influenced base substitutions only decreases to $n(1\text{-}10^{-5})$ after 10^2 years. Thus, the diploid state has the advantage to elongate the duration time of differentiated cells.

Estimation of the Positive Value of Biological Activity

The biological activity is first evaluated for the simplest case of RNA replicase and its estimation is then advanced to a free-living organism at the present time.

The RNA replicase

According to Cech [10], the RNA replicase self-reproduces in the following way, using the mononucleotides synthesized non-biologically. One type of RNA strand, a (+) strand, is synthesized on the template of complementary strand, a (-) strand, and conversely the (-) strand is synthesized on the template of a (+) strand, as shown in Figure 1. The template form of either a (+) or (-) strand is opened to be acceptable for mononucleotides, while the catalytic form takes a closed structure containing stem (intrastrand base-paired) regions. Mononucleotides are attached to the template form by the hydrogen bonds, and they are polymerized by forming the 3'-5' phosphodiester linkages under the action of the catalytic form.

The mononucleotide would have been a ribonucleoside-diphosphate or ribonucleoside-triphosphate just like in the present-day organism. The ribonucleoside-diphosphate contains one energy-rich bond $P \sim P$ and the ribonucleoside-triphosphate contains two energy-rich bonds $P \sim P \sim P$. The following release of Gibb's free energy ΔG is known in the decomposition of the energy-rich bonds [8].

ribonucleoside-P ~ P → ribonucleoside-P+P ΔG=6 Kcal/mole (2)

ribonucleoside-P ~ P ~ P → ribonucleoside-P+P ~ P ΔG=7 Kcal/mole (3)

P ~ P → P+P ΔG=6 Kcal/mole (4)

Upon the polymerization of mononucleotides, the energy is released not only from the above decomposition of energy-rich bonds but also from the formation of phosphodiester linkages, and the value of released energy is almost the same independently of the kinds of their bases. When the RNA replicase consists of (m+1) ribonucleotides, therefore, the energy of $m\Delta G$ or more is released upon its replication as the difference between the acquired energy and the stored energy, i.e., $Ea(M;N,S_N) - Es(N,S_N)$. The RNA replicase is further characterized by the special arrangement of four kinds of bases, adenosine (A), uracil (U), guanine (G) and cytosine(C), each contained in a ribonucleotide. Two hydrogen bonds can be formed between A and U, and three hydrogen bonds can be formed between G and C. By this intrinsic property, the special arrangement of nucleotide bases takes the catalytic form and the template form, alternatively, and it is transmitted from the template form to the replicated RNA replicase. The special arrangement of nucleotide bases means the lower entropy state in comparison with the random arrangement of them. Thus, the last term in the biological activity (1) of an RNA replicase consisting of (m+1) ribonucleotidesis expressed as

$$-TS_N=(m+1)RT\ln(1/4)=-(m+1)RT\ln 4 \quad (5)$$

where, R is the product of Avogadro number and Boltzmann constant, taking the value of 1.987 cal/deg. When temperature T is around 300°K, $RT\ln 4$ is less than*1* Kcal/mol. The number (m+1) is considered to have been 100 ~ 300 [9,10], and the biological activity $m\Delta G$ -(m+1) $RT\ln 4$ is certainly positive, even if the nucleoside-diphosphates are used as the

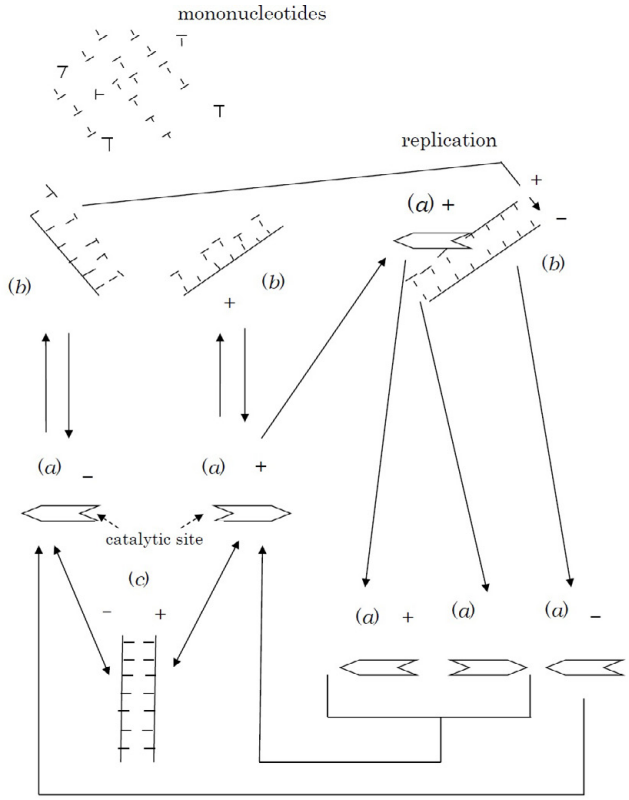

Figure 1: The outline of the behaviour of RNA replicase. The population of RNA replicases consists of (+) and (-) strands complementary to each other, and each strand alternatively takes two forms, (*a*) catalytic form and (*b*) template form, as shown on the left side of the figure. The catalytic form of (+) strand polymerizes the mononucleotides hydrogen-bounded to the template form of (-) strand, as shown at the upper part on the right side, and conversely the catalytic form of (-) strand also polymerizes the mononucleotides on the template form of (+) strand, although the latter process is omitted from the figure for simplicity. The replicases continue such replication, hardly falling into (*c*) the double stranded form of (+) and (-) strands, under the condition that the concentration of mononucleotides is sufficiently high.

mononucleotides. Thus, the presence of RNA replicases and the law of thermodynamics are not contradictory to each other. Although some nucleotide bases mismatched to the template RNA may have been inserted into the replica, this probability is estimated to be 10^{-4} per site from the strength of a hydrogen bond. Thus, such a mutant replicase only appears among 100 ~ 30 replicas.

An organism in the DNA-RNA-protein world

Although a free-living organism at the present time is much more complicated than the RNA replicase, the positive value of its biological activity can be estimated along the central dogma. The transcription and replication take place thermodynamically in the same way as the RNA replicase. The four kinds of ribonucleoside-triphosphates are also used in the transcription by the DNA-dependent RNA polymerase, and four kinds of deoxy forms are used in the replication of DNA by the DNA-dependent DNA polymerase, although the uracil-triphosphates are replaced by the deoxythimine-triphosphates. Upon the translation from mRNAs to proteins, each amino acid (*AA*) is first activated by adenosine-triphosphate (ATP) to form an energy-rich bond between the adenosine-monophosphate (AMP) and the *COO* group of the amino acid, and it is then transferred to the tRNA.

$$\text{AA+adenosine-P} \sim \text{P} \sim \text{P} \rightarrow \text{AA} \sim \text{AMP+P} \sim \text{P} \quad (6)$$

$$\text{AA} \sim \text{AMP+tRNA} \rightarrow \text{AA} \sim \text{tRNA+AMP} \quad (7)$$

The amino acid residue carried by each tRNA is combined to grow a peptide chain by forming a peptide bond on the ribosome, according to the triplet bases (codon) of the mRNA under translation.

$$\text{AA} \sim \text{tRNA+polypeptide chain of p amino acid residues} \rightarrow \text{tRNA+polypeptide chain of (p+1) amino acid residues} \quad (8)$$

Because twenty kinds of *L*-type amino acids are used in this translation, the special arrangement of q amino acid residues in the resultant polypeptide means the lower entropy state of $-qR\ln 20$ in comparison with the random arrangement of them. Such lower entropy is also sufficiently compensated for by the entropy production due to the heat released from the decomposition of energy-rich bonds in the processes Eq. (6-8) and from the formation of peptide bonds.

These proteins translated from respective genes exhibit various biological functions depending on their amino acid sequences, that is, they serve to acquire the material and energy sources from the outside, and to catalyze specifically the synthetic pathways converting the material sources into nucleotides, amino acids and other biomolecules for the construction of cell structure as well as the degradation pathways of organic compounds. In the degradation pathways represented by the glycolysis and respiration, the AMPs are converted into the ATPs, and the energy of ATPs is used in the above synthetic path ways as well as in the conversion of other kinds of ribonucleoside-mono and -diphosphates into the ribonucleoside-triphosphates. The coupling of the synthetic pathways with the degradation pathways through ATPs strongly suggests that the entropy production by the heat released from these pathways compensates for the lower entropy of taking the special metabolic pathways, as a whole. In the multicellular eukaryotes, the systematization of cell differentiation takes place by cell type specific ligands and receptors and by the intracellular signal transduction mainly of phosphorylation from the ligand-accepted receptor to transcriptional regulators. Upon this communication between cells, the energy-rich bonds in GTPs are decomposed and the released heat compensates for the entropy reduction of systematization between differentiated cells, thermodynamically. Thus, the positive value of biological activity may hold for the multicellular organism as well as the unicellular one.

Application of the Concept of Biological Activity

Although the absolute value of biological activity cannot be easily obtained for the present-day organism, its changes by the variation in the internal variables (N,S_N) and in the mode of transcription become useful measures to analyze the biological phenomena. These two examples will be presented in this section.

Evolution of organisms in the DNA-RNA-protein world

Since the publication of "The origin of species" by Darwin [6], Darwinian evolution by the nucleotide base change and selection has been widely accepted in biology [15,16]. This is also the process to maintain the lower entropy state of survived organisms, although it is not noticed in physics as well as in biology. Moreover, recent gene and genome sequencing reveals that the repertoire of protein functions has been expanded by gene duplication and by the succeeding changes in the counterpart of duplicated genes [17-20], suggesting the larger scale of evolution than Darwinian evolution. The biological activity becomes a useful measure for the theoretical formulation of such large-

scale evolution. This will be simply shown for the monoploid organism. First, gene duplication enlarges the genome size to $N+\Delta N$. This raises the stored energy to $E_s(N+\Delta N,S_N)$, while the value of acquired energy $E_a(M;N+\Delta N,S_N)$ is almost the same as that of $E_a(M;N,S_N)$. Thus, the biological activity of the variant having experienced gene duplication first becomes lower than that of the original style organism, especially at the stage when the counterpart of duplicated genes loses the original function by its changes. However, such variants are not necessarily compelled to extinction but can exist as minor members in the population. Moreover, new function(s) can arise from such a changing gene. If the product(s) of such new gene(s) become suitable for the variant to acquire a new material and energy source L and/or to move to a new area, the acquired energy $E_a(L;N+\Delta N,S_{N+\Delta N})$ turns to increase, overcoming the increase in systematization from S_N to $S_{N+\Delta N}$ as well as in stored energy $E_s(N+\Delta N,S_{N+\Delta N})$. Thus, a new style organism appears with the recovered biological activity $E_a(L;N+\Delta N,S_{N+\Delta N})$ $-E_s(N+\Delta N,S_{N+\Delta N})$ $-TS_{N+\Delta N}$. This large-scale evolution explains the divergence of new and original styles of organisms [21,22], without the geographical isolation and/or climate change required for the generation of new species in Darwinian evolution. The mathematical description of the large-scale evolution is shown more strictly in Appendix, where the distinction between the large-scale evolution and Darwinian evolution is more clearly represented.

The large-scale evolution of diploid organisms is accompanied by the additional process to establish the new style organisms as those carrying the new gene(s) homologously [23], and this process explains the punctuated mode of explosive divergence of new morphological organisms suggested from the paleontology and the reconstructed tree of phylogeny [24,25]. The divergence of organisms in the same area also gives a theoretical basis for forming an ecological system, for which the concept of biological activity also holds in a wider sense [26].

The developmental process of a multicellular organism

Any multicellular organism grows from a unicellular spore or fertilized egg by increasing the number of cells to differentiate into distinctive types. The biological activity changes during such development, although the biological activity throughout the lifetime of an organism is considered for the investigation of evolution. For its estimation, the change in biological activity will be expanded in terms of the number of differentiated cells near the initiation of development.

In the case when the cells differentiate into two types, the change in biological activity $\Delta BA(n_1, n_2)$is expressed in the following form by using the number n_1 of type 1 of cells and the number n_2 of type 2 of cells.

$$\Delta BA\ (n_1,n_2)=-a_1n_1-a_2n_2+2b_{12}n_1n_2 \qquad (9)$$

where a_1, a_2 and b_{12} are positive coefficients. The first and second terms on the right side of Eq. (9) indicate the increase in stored energy and systematization concerning types 1 and 2 of cells, respectively, by the beginning of transcription, translation and replication, and the third term is the energy acquired from the outside by the cooperative action of differentiated cells. According to the knowledge of the cell cycle, the cell multiplication and the cell differentiation do not proceed simultaneously. Thus, the type 1 of cells first increase their number until they go over the threshold of $n_1=a_2/2b_{12}$. In this first step, the energy is decreased with the amount of $a_1=a_{11}$ per cell by transcribing the genes for the multiplication and exhibition of type 1 character. Then, the first type cells further transcribe the second group of genes to induce type 2 of cells and to communicate with the latter. By this gene expression, the coefficient a_1 of the first type cell is further increased by a_{12} per cell. Under the increased value of $a_1=a_{11}+a_{12}$, the type 2 cells appear with $a_2=a_{22}$, and increase their number n_2 to enter the region of $\Delta BA(n_1,n_2)>0$ by the cooperative action of the two types of cells, as shown in Figure 2.

In the case when the cells differentiate into three types, the change in biological activity $\Delta BA(n_1,n_2,n_3)$ is expressed in the following way by using the number n_3 of the third type cells as well as $n_1$and n_2.

$$\Delta BA(n_1,n_2,n_3)= -a_1n_1-a_2n_2-a_3n_3+2b_{12}n_1n_2+2b_{13}n_1n_3+2b_{23}n_2n_3 \qquad (10)$$

In this case, the first and second types of cells appear in the same way as in the preceding case of Eq. (9). Then, the first and second types of cells further transcribe the third group of genes to induce the third type of cells and to communicate with them. By this gene expression, the coefficients a_1 and a_2 further increase to $a_{11}+a_{12}+a_{13}$ and $a_{22}+a_{23}$, respectively. Under this situation, the third type cells appear and increase their number to enter the region of $\Delta BA(n_1,n_2,n_3)>0$, as shown in Figure 3. This pathway is consistent with the fundamental pattern of development observed in higher animals of *Triploblastica*; the first type cells correspond to those forming the ectoderm, the second to the endoderm and the third to the mesoderm. In the higher plants, the roots first appear, then the leaves appear and the stem is formed.

The above analysis reveals the followings. First, the change BA in biological activity takes the negative value until the cooperative action of differentiated cells increases the acquired energy. This explains the fact that the parent must endow the spore or egg with material and energy sources. Consuming the endowed sources, the cells in development increase their stored energy and systematization to exhibit the cooperative action, during which the endowed energy E_{en} plus ΔBA retains positive values consistently with the law of thermodynamics. Second, more energy and materials are needed for the development of higher hierarchy of cell differentiation. Moreover, the energy required for development becomes larger in the diploid state (fertilized egg) than in the monoploid state (spore), because the coefficients a_i' s (i=1,2,3) are larger for the transcription and replication of diploid state. On the other hand, the diploid state assures the longer duration time of differentiated cells, as indicated already. The lower eukaryotes alternating the monoploid generation with the diploid one illustrate the evolutionary route along which monoploid organisms have climbed up to diploid organisms with the advancement in sexual organs by the repetition of generating new genes from gene duplication [25]. The biological activity during the development is similar to but different from the 'potential function', which is proposed in the application of catastrophe theory for the morphogenesis [27]. Such potential function should be considered about the interaction between the surfaces of differentiated cells and is considerably different between animals and plants.

Conclusion and Discussion

The genes responsible for the lower entropy state of an organism are maintained and further improved by the self-reproduction and selection. The thermodynamic quantity of biological activity representing the state of such an organism not only resolves the paradox of Maxwell's demon but also leads to the comprehensive understanding of biological phenomena, complementally to the knowledge of molecular biology. In particular, the large-scale evolution indicates that self-reproducing cells containing the extra polynucleotides to be replicated have the potentiality to produce new genes as long as their biological activity is positive. This mechanism of self-reproducing cells

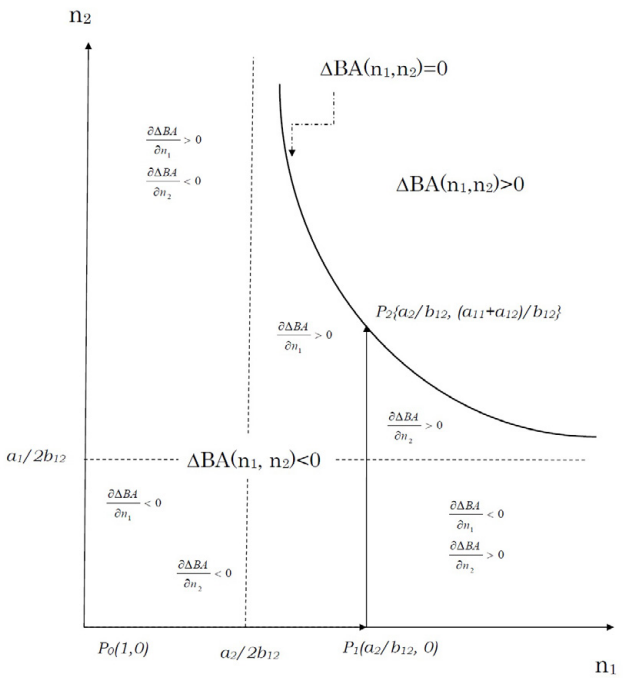

Figure 2: The change in biological activity during the development of differentiation into two types of cells. According to Eq. (9), the curve of $\Delta BA(n_1,n_2)=0$ is plotted in the (n_1,n_2) plane under the positive values of $a_1(=a_{11}+a_{22})$, $a_2(=a_{22})$ and b_{12}. The region of $\Delta BA(n_1,n_2)>0$ is separated from the region of $\Delta BA(n_1,n_2)<0$, which contains the unicellular state (n_1=1 and n_2=0), by the hyperbolic curve $\Delta BA(n_1,n_2)=0$. The changes in $\Delta BA(n_1,n_2)$ due to the increase in n_1 and n_2 are different between the four regions divided by two dotted lines, $n_1=a_2/2b_{12}$ *and* $n_2=a_1/2b_{12}$ as shown in the figure. Starting from the unicellular state $P_0(1,0)$, therefore, the process of cell differentiation must proceed in the region of $\Delta BA(n_1,n_2)<0$ until it reaches the region of $\Delta BA(n_1,n_2)>0$. As discussed in the text, the most plausible pathway first progresses on the n_1 axis with the coefficient $a_1=a_{11}$ until the value of n_1 becomes larger than $a_2/2b_{12}$, and then turns to take the path along the n_2 axis, where the coefficient a_1 is increased to $a_{11}+a_{12}$ for inducing the second type of cells and communicating with them. This pathway of development is shown by $P_0(1,0) \rightarrow P_1(a_2/b_{12}, 0) \rightarrow P_2\{a_2/b_{12},(a_{11}+a_{12})/b_{12}\}$.

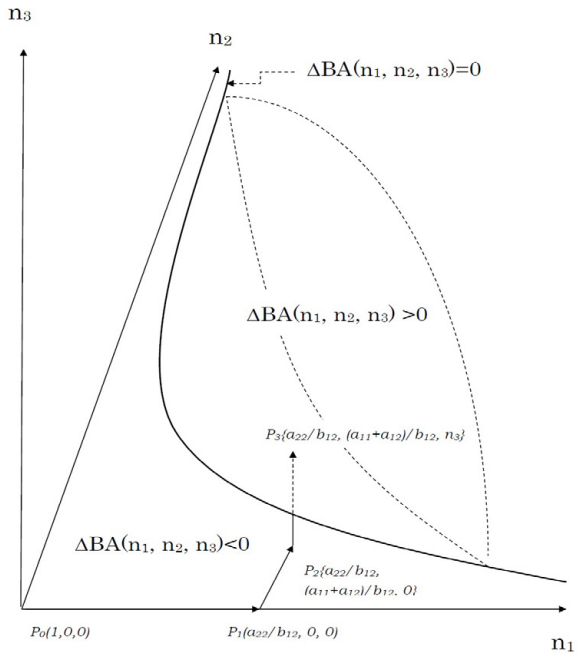

Figure 3: The change in biological activity during the development of differentiation into three types of cells. Under the positive values of coefficients $a_1(=a_{11}+a_{12}+a_{13})$, $a_2(=a_{22}+a_{23})$ *and* a_3, $\Delta BA(n_1,n_2,n_3)=0$ of Eq. (10) plotted in the (n_1,n_2,n_3) space is the hyperboloid of a sheet, whose section is shown by a dotted ring, and $\Delta BA(n_1,n_2,n_3)>0$ holds in the region where the values of n_1, n_2 and n_3 are larger than those on the sheet. The developmental pattern of higher animals and plants indicates the following pathway to enter the region of $BA(n_1,n_2,n_3)>0$ from the unicellular state $P_0(1,0,0)$. Fist, only n_1 increases to reach the point $P_1(a_{22}/b_{12},0,0)$ on the n_1 axis, and then n_2 value increases to reach the point $P_2\{a_{22}/b_{12},(a_{11}+a_{12})/b_{12}, 0\}$ in the same way as the pathway in Figure 2. However, this point P_2 is still in the region $\Delta BA(n_1,n_2,n_3)<0$, because a_1 and a_2 are increased by a_{13} *and* a_{23} respectively to induce the third type cells. Under the condition of $a_1=a_{11}+a_{12}+a_{13}$ *and* $a_2=a_{22}+a_{23}$, the third type of cells appear, increasing their number n_3 to reach the point $P_2\{a_{22}/b_{12},(a_{11}+a_{12})/b_{12},n_3\}$on the sheet. Thus, the energy required for the development into three types of cells becomes larger than the energy for the differentiation into two types of cells, at least by the third step.

suggests a possible way about how the RNA replicases have evolved to the present-day organisms.

In the RNA world where the RNA replicases lived, various organic compounds including amino acids and nucleotides would have been synthesized non-biologically. Some satellite variants of RNA replicases would have catalyzed the polymerization of amino acids [28]. Even if the polypeptides thus generated each consist of several amino acid residues, the random sequences of five amino acid residues amount to 20^5 kinds of polypeptides, which sufficiently cover the active centres of enzyme proteins in a free-living organism at the present time [29]. By the catalytic actions of these polypeptides, it is plausible that primitive cells are formed and they self-reproduce by the cell division due to the antagonistic balance of lipid synthesis and cell wall construction [30]. The ancient origin of cell wall before the appearance of translation system is suggested from the prokaryotes such as *Escherichia coli* where proteins catalyze the polymerization of *L*-and *D*-types of amino acids to form the cell wall elements as a molecular fossil. If the cell contains the polypeptides with RNA polymerase activity stronger than the RNA replicase, the RNAs in the cell are allowed to change apart from the functional constraint of RNA replicase activity. In such proto-cells, the primitive translation system would have appeared through the interference among different kinds of increased RNAs, even if the biological activity was transiently lowered by the interference. The systematic force of selection must have then established the protein synthesis based on codon usage and converged other RNAs to the RNA genes of proteins useful for enhancing the self-reproduction of cells. The deoxidization of RNAs would have also occurred in such self-reproducing cells that attempted the glycolysis to acquire the more energy. The DNAs generated by the protons released from the glycolysis would have been first rubbish, but gradually induced the genes of proteins necessary for their transcription and replication as the derivatives of RNA polymerase gene [31]. After the parallel usage of DNA genes and RNA genes, the decisive turning point to the DNA-RNA-protein world would have been the beginning of equi-partition of replicated DNAs into daughter cells by the direct or indirect attachment to the cell membrane, bringing about the divergence of prokaryotes and eukaryotes. The DNA genes in the prokaryote are fused to a single circular molecule suitable for the cell division by the direct attachment to the membrane upon its replication while the DNA genes in the eukaryote get together to the plural number of chromosomes and two

sets of them arising from replication are equi-partitioned into daughter cells by attaching to the membrane through contractive microtubles. In fact, the analysis of nucleotide base changes in ribosomal RNAs assures that all organisms in the DNA-RNA-protein world are traced back to the ancient divergence of prokaryotes and eukaryotes [32,33]. This innovation not only raises the accuracy of cell division but also makes it possible to evolve synthetic pathways extensively on the basis of the stable DNA genes, overcoming the decrease in organic compounds synthesized non-biologically. Moreover, some of the eukaryotes further evolve to multicellularity and diploid state after the acquirement of the mitochondria and chloroplasts as the endosymbionts of O_2-respiratory and photosynthetic eubacteria. Meanwhile, the original style organisms in the RNA-protein world have turned their strategy to utilize the nucleotides and amino acids as well as the ribosome in the organisms evolved to the DNA-RNA-protein world and survive as RNA viruses, although the RNA replicases would have no means to survive. In this connection, it should be noted that the biological activity is degenerate, that is, almost the same degree of biological activity can be attained by either a small genome, low systematization and a small amount of acquired energy or a large genome, high systematization and a large amount of acquired energy. The mathematical formulation of the above innovation from the proto-cells to the present-day organisms is also possible as the extension of the formulation given in Appendix. Thus, circumstantial evidences wait for experimental approaches to the construction of self-reproducing proto-cells under the pre-biotic conditions.

References

1. Maxwell JC (1871) Theory of heat. For the detailed description of Maxwell's demon. J Appl Phys 22: 334-337.
2. Szilard L (1926) On the entropy reduction in a thermodynamic system in the case of the intervention of intelligent beings. Z Physik 53: 840-856.
3. Brillouin L (1951) Maxwell's demon cannot operate: information and entropy. I J Appl Phys 22: 334-337.
4. Prigogine I (1967) Introduction to thermodynamics of irreversible process (3rd edn.). Interscience Publishers, a Division of John Wiley & Sons, New York, London, Sydney.
5. Eigen M, Schuster P (1977) The hypercycle. A principle of natural self-organization. Part A: Emergence of the hypercycle. Naturwissenschaften 64: 541-565.
6. Darwin C (1859) The origin of species. John Murry, London.
7. von Neumann J (1966) Theory of self-reproducing automata. The University of Illinois Press, Chicago IL, USA.
8. Watson JD, Hopkins NH, Roberts JW, Steitz JA et al. (1987) Molecular biology of the gene. (4th edn.) The Benjamine/Cumming Publisher Company, California, USA.
9. Zaug AJ, Cech TR (1986) The intervening sequence RNA of Tetrahymena is an enzyme. Science 231: 470-475.
10. Cech TR (1986) A model for the RNA-catalyzed replication of RNA. Proc Natl Acad Sci USA 83: 4360-4363.
11. Otsuka J, Nozawa Y (1998) Self-reproducing system can behave as Maxwell's demon: Theoretical illustration under prebiotic conditions. J Theor Biol 194: 205-221.
12. Kimura M (1980) A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. J Mol Evol 16: 111-120.
13. Otsuka J, Nakano T, Terai G (1997) A theoretical study on the nucleotide changes under a definite functional constraint of forming stable base-pairs in the stem regions of ribosomal RNAs; its application to the phylogeny of eukaryotes. J Theor Biol 184: 171-186.
14. Wheeler DL, Church DM, Edgar R, Federhen S, Helmberg W et al. (2004) National Center of Biotechnology Information. Nucl Acid Res 32.
15. Fisher RA (1930) The general theory of natural selection. Oxford University Press, London.
16. Wright S (1949) Adaptation and selection. In: Genetics, palaeontology and evolution pp: 365-389.
17. Ingram VM (1963) The hemoglobin in genetics and evolution. Columbia Press, NY, USA.
18. Ohno S (1970) Evolution by gene duplication. Spring-Verlag, Berlin.
19. Gilbert W (1978) Why genes in pieces? Nature 271: 501.
20. Ferris SD, Whitt GS (1979) Evolution of the differential regulation of duplicated genes after polyploidization. J Mol Evol 12: 267-317.
21. Otsuka J (2005) A theoretical scheme for the large-scale evolution of organisms towards a higher order of organization and diversity. In: Recent research developments in experimental and theoretical biology 1: 93-122.
22. Otsuka J (2008) A theoretical approach to the large-scale evolution of multicellularity and cell differentiation. J Theor Biol 255: 129-136.
23. Otsuka J (2011) The large-scale evolution by generating new genes from gene duplication; similarity and difference between monoploid and diploid organisms. J Theor Biol 278: 120-126.
24. Otsuka J (2011) A theoretical scheme of the large-scale evolution by generating new genes from gene duplication. In: Friedberg F (ed.) Gene duplication. In Tech, Croatia pp: 3-26.
25. Otsuka J (2014) The large-scale evolution by generating new genes from gene duplication and the divergence pattern of organisms. In: Advances in genetics research 11: 85-116.
26. Otsuka J (2004) A theoretical characterization of ecological systems by circular flow of materials. Ecological Complexity 1: 237-252.
27. Thom R (1983) Mathematical models of morphogenesis. Translated by Brooks WM, Ellis Horwood Limited, a division of John Wiley&Sons.
28. Weiner AM, Maizels N (1987) tRNA-like structures tag the 3' ends of genomic RNA molecules for replication: implications for the origin of protein synthesis. Proc Natl Acad Sci U S A 84: 7383-7387.
29. Fukuchi S, Otsuka J (1996) Evolution of the self-reproducing system to the biosynthesis of the membrane: an approach from the amino acid sequence similarity in proteins. J Theor Biol 182: 117-136.
30. Otsuka J (2001) Cell division potentiality arising from the antagonistic balance of lipid synthesis and cell wall construction. Research Communication in Biochemistry, Cell & Mol Biol 5: 391-416.
31. Otsuka J, Kikuchi N, Kojima S (1999) Similarity relations of DNA and RNA polymerases investigated by the principal component analysis of amino acid sequences. Biochim Biophys Acta 1434: 221-247.
32. Woese CR (1987) Bacterial evolution. Microbiol Rev 51: 221-271.
33. Sugaya N, Otsuka J (2002) The lineage-specific base-pair contents in the stem regions of ribosomal RNAs and their influence on the estimation of evolutionary distances. J Mol Evol 55: 584-594.

Citation: Otsuka J (2017) The Concept of Biological Activity and Its Application to Biological Phenomena. J Phys Chem Biophys 7: 235. doi: 10.4172/2161-0398.1000235

Appendix

Mathematical formulation of Darwinian evolution and the large-scale evolution

The set of internal variables (N,S_N) of a monoploid organism will be simply denoted as a single variable x, unless the description of changes in genome size is necessary. Because the organism is generally characterized by self-reproduction, mutation and death, the population of such organisms taking a common material and energy source M from the environment consists of various variants, and the number $n_{xi}(t)$ of i-th variants x_i obeys the following time change equation.

$$\frac{d}{dt}n_{xi}(t)=\{Q_{xi}(t)R(M;x_i)-D(x_i)\}n_{xi}(t)+\sum_{j(\neq i)}q_{xi,xj}(t)R(M;x_j)n_{xj}(t) \quad \text{(A1)}$$

where the self-reproducing rate and death rate of the variant x_i are denoted by $R(M;xi)$ and $D(xi)$, respectively. The apparent decrease factor $Q_{xi}(t)$ for the self-reproducing rate of the variant x_i means the mutation of variant xi to other kinds of variants and is related with the mutation term $q_{xj,xi}(t)$ in the following way:

$$Q_{xi}(t)=1-\sum_{j(\neq i)}q_{xj,xi}(t) \quad \text{(A2)}$$

For the investigation of the population behaviour, it becomes transparent to transform Eq. (A1) into two types of equations; one concerning the total number $B(t)$ of all kinds of variants defined by

$$B(t)=\sum_i n_{xi}(t) \quad \text{(A3)}$$

and another concerning the fraction $f_{xi}(t)$ of the variants x_i defined by $n_{xi}(t)/B(t)$. By simple calculation, these equations are obtained from Eq. (A1) to the following forms, respectively, using the relation (A2):

$$\frac{d}{dt}B(t)=W_{av}(M;t)B(t) \quad \text{(A4)}$$

and

$$\frac{d}{dt}f_{xi}(t)=\{W(M;x_i)-W_{av}(M;t)\}f_{xi}(t)+\sum_j q_{xi,xj}(t)R(M;x_j)f_{xj}(t) \quad \text{(A5)}$$

where $q_{xi,xi}(t)$ is defined by $1–Qxi(t)$. The increase rate $W(M;x_i)$ of the variant x_i and the average increase rate $W_{av}(M;t)$ of the organisms in the population are defined by

$$W(M;x_i)\equiv R(M;x_i)-D(x_i) \quad \text{(A6)}$$

and

$$W_{av}(M;t)\equiv\sum_i W(M;x_i)f_{xi}(t) \quad \text{(A7)}$$

respectively.

Darwinian evolution corresponds to the following approximate evaluation of Eq. (A5) with respect to the change in the fraction of variants. If the increase rate $W(M;xi)$of occasionally arisen mutant x_i is larger than the average increase rate of the population, that is, $W(M;xi)-Wav(M;t)>0$, the fraction $f_{xi}(t)$of the mutants x_i increases with time according to the first term on the right side of Eq. (A5). This raises the average increase rate $Wav(M;t)$, resulting in the increase in the total number $B(t)$ of organisms according to Eq. (A4), although this increase is ultimately stopped by the decrease in available material and energy source M. On the other hand, the fraction $f_{xi}(t)$ decreases when $W(M;xi)-Wav(M;t)<0$. Thus, the organisms taking a common material and energy

source *M*are elaborated by mutation and selection, and most of them finally reach the ones with optimum increase rate, each characterized by x_{opt}. In such Darwinian evolution, the mutants are mostly due to the nucleotide base substitutions in the existing genes, which are mentioned in the text. Thus, Darwinian evolution gives the theoretical reason why the lower entropy state of the special nucleotide base sequence of genes is maintained in the survived organism, although this evolution has been only focused on the generation of new species as the adaptation to a geographically isolated region and/or to climate change in biology.

For the mathematical formulation of the large-scale evolution by generating new gene(s) from gene duplication, it is necessary to consider the fraction of variants more accurately. For this purpose, Eq. (A5) will be formally integrated with respect to time *t*, i. e.,

$$f_{xi}(t)=\exp\left[\int_0^t\{W(M;x_i)-W_{av}(M;\tau)\}\,d\tau\right]\left[\int_0^t\left\{q_{xi,xj}(\tau)R(M;x_j)\right\}f_{xj}(\tau)\times\exp\left[-\int_0^t\{W(M;x_i)-W_{av}(M;\tau')\}\,d\tau'\right]d\tau+f_{xi}(0)\right] \quad \text{(A8)}$$

In this expression of the fraction $f_{xi}(t)$ of variants x_i, the average increase rate $W_{av}(M;t)$ is approximated to be $W(M;x_{opt})$ and the fractions of other variants except for x_{opt} are neglected on the right side of Eq. (A8) after the organisms characterized by x_{opt} have been dominant in the population. Then, it is shown that the fraction $f_{xi}(t)$ of the variants x_i with the lower increase rate $W(M;x_i)$ is not necessarily zero but can be persistently present with the following relation to the fraction f_{xopt} of dominant organisms.

$$f_{xi}=\frac{q_{xi,xopt}R(M;x_{opt})}{W(M;x_{opt})-W(M;x_i)}\,f_{xopt} \quad \text{(A9)}$$

where $q_{xi,xopt}$ is the mutation rate defined by an average of mutation terms during a sufficiently long time *t*, i. e.,

$$q_{xi,xopt}=\frac{1}{t}\int_0^t q_{xi,xopt}(\tau)d\tau \quad \text{(A10)}$$

Among such satellite variants, the variant having experienced the gene duplication is especially notable in the sense that it has the potential to generate new gene(s) from the counterpart of duplicated genes, although the gene duplication occurs less frequently than the nucleotide base substitutions. To follow such large-scale evolution, the satellite variant characterized by the set of internal variables $(N+\Delta N,S_N)$will be denoted as x_{in} and the original style of a dominant organism characterized by x_{opt} will be re-denoted by x_o, hereafter. A new style of an organism $(N+\Delta N,S_{N+\Delta N})$, whose internal variables will be simply denoted as *y*, can arise from the intermediate variant x_{in} by generating new gene(s) from the spare gene in ΔN and by advancing the systematization to $S_{N+\Delta N}$. Denoting this probability by $q_{y,xin}$, the transition probability $P(y\leftarrow x_{in}\leftarrow x_o)$ from the original style organism to a new style organism is expressed by the following form, according to (A9).

$$P(y\leftarrow x_{in}\leftarrow x_o)=\frac{q_{y,xin}q_{xin,xo}R(M;x_o)}{W(M;x_o)-W(M;x_{in})} \quad \text{(A11)}$$

If the new gene(s) serve to utilize the material and energy sources *M* more efficiently, the new style organisms compel the original style organisms to extinction. If the new gene(s) give the new style organism the ability of utilizing new material and energy source *L* other than *M* and/or of moving to a new area, on the contrary, the

new style organisms form a new population, apart from the population of the original style organisms. In this new population, the fraction $f_{yk}(t)$ of the new style organisms y_k obeys the following equation.

$$\frac{d}{dt} f_{yk}(t) = \left\{ W(L; y_k) - \overline{W}(t) \right\} f_{yk}(t) + \sum_k q_{yk,yk}(t) R(L; y_l) f_{yl}(t) \qquad \text{(A12)}$$

where

$$\overline{W}(t) \equiv \sum_i W(M; x_i) f_{xi}(t) + \sum_k W(L; y_k) f_{yk}(t) \qquad \text{(A13)}$$

The ability of new gene(s) is elaborated to raise the increase rate in the new population by Darwinian evolution. The total number $B(t)$ of organisms is re-defined as

$$B(t) \equiv \sum_i n_{xi}(t) + \sum_i n_{yk}(t) \qquad \text{(A14)}$$

and obeys the following equation.

$$\frac{d}{dt} B(t) \equiv \overline{W}(t) B(t) \qquad \text{(A15)}$$

This divergence of new and original styles of organisms means that the total number of organisms increases by taking different material and energy sources and/or by living in different areas. After the new style organisms y_{opt} are established, the arising of the same style of new organisms from the original style organisms is ceased, because any primitive organism y newly arising from the original style organisms tends to extinction by the struggle of existence with the new style organisms y_{opt}.