The Sixteenth International Symposium on Artificial Life and Robotics 2011 (AROB 16th '11),
B-Con Plaza, Beppu,Oita, Japan, January 27-29, 2011

The inevitability of the bio-molecules: five nitrogenous bases and twenty amino acids

K.Hashimoto, H.Inoue and K.Naitoh
Waseda University, Faculty of Science and Engineering, 3-4-1 Ookubo, Shinjuku, Tokyo, 169-8555 Japan
(Tel : 81-3-5286-3265)
(Email address:k-naito@waseda.jp)

Abstract: Living beings meta-stable between unstable and neutral-stable conditions use only five types of nitrogenous bases and twenty amino acids selected naturally. The unified momentum theory derived at the triple point of the deterministic Boltzmann equation, the stochastic Langevin equation, and the indeterminacy principle [K. naitoh, Artificial Life Robotics, 2010] reveals the reason why the molecular weights of the twenty types of amino acids show a threefold variation between 240 of cysteine as the maximum and 75 of glycine as the minimum, whereas that of purines and pyrimidines among nitrogenous bases varies by only about 1.5 times, although the variation principle for the energy conservation law cannot do so. Let us take a higher order of the Taylor series for the unified momentum equation describing the deformation motions of biological molecules. Even-numbered terms such as the second and fourth ones show no other quasi-stable size ratios. However, odd-numbered terms result in other quasi-stable ratios. The threefold variation of amino acids will come from the third term of the Taylor series.

***Keywords*: Asymmetry, Cyto-fluid dynamics**

I. INTRODUCTION

Many types of amino acids and bases are possible in the artificial pools of molecules, but living beings meta-stable between unstable and neutral-stable conditions use only five types of nitrogenous bases and twenty amino acids selected naturally. [1-6]

Life can exist only when fully surrounded by water molecules. Thus, the meta-stability of biological molecules such as nitrogenous bases, nucleic acids, amino acids, and proteins in living beings should be analyzed by considering the flows of water molecules.

Our previous model based on traditional fluid dynamics [7-9] has qualitatively revealed the reason why living beings employ only five nitrogenous bases, i.e., two purines of A and G and three pyrimidines of T, C, and U, and also why purines and pyrimidines have a size ratio of about 1.5. However, the model was with some gratuitous assumptions and also analyzed by the first order of accuracy, which led only to qualitative results. Consequently, a qualitative analysis did not reveal the reason why the molecular weights of the twenty types of amino acids show a threefold variation between 240 of cysteine as the maximum and 75 of glycine as the minimum and also why proteins are more various.

In this report, a unified momentum theory is extended from our previous model as a result of statistical mechanics [9] and the theory of adding mass [10]. A higher order of analysis based on the momentum conservation law reveals the size varieties of the five bases and twenty amino acids. The present report reveals several mysteries underlying the bio-molecular systems.

II. CYTO-FLUID DYNAMICS THE ORYBASED ON the γ-ε EQUATION

First, we define a parcel as a flexible spheroid having two long and short radii of $a(t)$ and $b(t)$ dependent on time t and constructed of a bio-molecule such as a base or an amino acid that is surrounded by water molecules. (A parcel consists of a bio-molecule such as nitrogenous base or amino acid, water molecules hydrated with the bio-molecule, and the immerse mass due to adding mass effect while impulsive start and stop of deformations occur [8, 9].)

The parcel size is proportional to that of biological molecules such as nitrogenous base or amino acid.

The parcel becomes a sphere of the radius $r_d (= [ab^2]^{1/3})$ under an equilibrium condition. The deformation rate $\gamma(t)$ is defined as $a(t)/b(t)$, while a sphere without deformation corresponds to $\gamma = 1$.

We assume that the flow field inside the parcel is the potential flow.

We derive a theory for describing the deformation and motions of two connected spheroid parcels having two equilibrium radii of r_{d1} and r_{d2} and two deformation rates of γ_k $[k = 1, 2]$, while the size ratio of the two parcels is defined by $\varepsilon = r_{d1} / r_{d2}$. We model the one-dimensional relative motion between the two parcels, nonlinear convections inside the parcels, the interfacial force at the parcel surface due to forces interacting between molecules, and collisions with water molecules outside the parcels. The interfacial force is evaluated in the form of σ / r^m where m and σ are constants and r is the curvature of parcel surface. [7] Several types of forces such as van der Waals force, coulomb force, and surface tension can be explained by varying m. Here, the relation m = 1 implies the surface tension of liquid. The mean density of the parcels is ρ_L.

Moreover, the scale for averaging, i.e., the minimum scale representing the phenomenon, will be smaller than that in continuum mechanics and will be between the atomic scale and the size of the bio-molecule. Thus, this small averaging window applied to the Boltzmann equation [9] leads to a weak indeterminacy of physical quantities such as deformation and density because of molecular discontinuity.

Next, we derive the momentum equation describing the relation between a dimensionless deformation rate $\gamma_k (\equiv a_k / b_k \ [k = 1, 2])$ of each parcel dependent on dimensionless time $\bar{t}_k = \sqrt{\dfrac{8\sigma}{\rho_L r_{dk}^{2+m}}}\, t \quad [k = 1, 2]$ and the size ratio of the two parcels of $\varepsilon = r_{d1} / r_{d2}$. The stochastic governing equation having indeterminacy can be described as

$$\begin{aligned}\frac{d^2}{d\bar{t}_i^{\,2}}\gamma_i = \{ & [(-\varepsilon - \varepsilon^4 + \frac{2}{3}\varepsilon E_{0j}\gamma_j^{-1/3})B_{0i} + \frac{2}{9}\varepsilon^{2+m}E_{0i}\varepsilon\gamma_i^{-4/3}](\frac{d}{d\bar{t}_i}\gamma_i)^2 \\ & + [\frac{2}{3}\varepsilon^{2+m}E_{0i}\gamma_j^{-1/3}B_{0j} - \frac{2}{9}\varepsilon^{2+m}E_{0i}\gamma_j^{-4/3}](\frac{d}{d\bar{t}_j}\gamma_j)^2 \\ & + (-\varepsilon - \varepsilon^4 + \frac{2}{3}\varepsilon E_{0j}\gamma_j^{-1/3})C_{0i}\ \gamma_i^{\frac{5}{3}-\frac{2}{3}m} + \frac{2}{3}\varepsilon^{2+m}E_{0i}\gamma_j^{-1/3}C_{0j}\ \gamma_j^{\frac{5}{3}-\frac{2}{3}m}\} / Det \\ & + \delta_{st}\end{aligned}$$

$[\text{for } i = 1, 2.\ j = 1, 2. i \neq j]$ (1)

with

$$Det = -\varepsilon - \varepsilon^4 + \frac{2}{3}\varepsilon^4 E_{0i}\gamma_i^{-1/3} + \frac{2}{3}\varepsilon\ E_{0j}\gamma_j^{-1/3},\ B_{0k} = \frac{1}{3\gamma_k}\frac{\gamma_k^2 - 2}{\gamma_k^2 - 1/2},$$

$$C_{0k} = \frac{3}{8}\frac{2\gamma_k^{2m} - 1/\gamma_k^m - \gamma_k^m}{\gamma_k^2 - 1/2}, \text{ and } E_{0k} = 3\frac{\gamma_k^{7/3}}{\gamma_k^2 - 1/2} [\text{for } k = 1, 2]$$

where the parameter δ_{st} denotes random fluctuation.

The long derivation of Eq. (1) is in Ref. 6 confirmed by the referees, although only the stochastic term δ_{st} is not in Ref. 6.

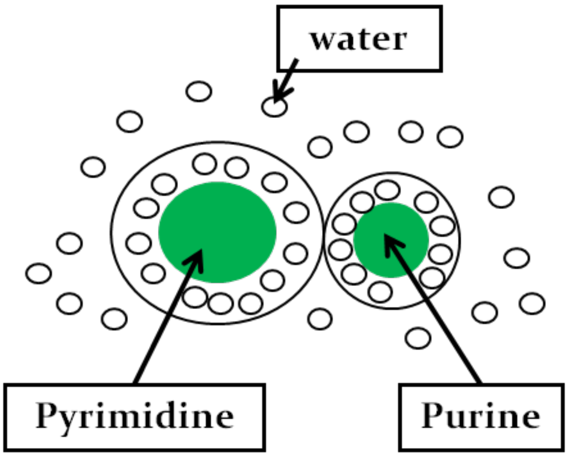

Fig.1. Two parcels connecting

III. FIRST ORDER OF ANALYSIS [6]

We then define the deviation from a sphere as y_i, which is equal to $\gamma_i - 1$. Taking the first order of approximation in the Taylor series leads to

$$\begin{aligned}\frac{d^2 y_i}{d\bar{t}_i^{\,2}} = [& -\frac{2}{3}(3 - \varepsilon^3 - 2\varepsilon^{2+m})\left(\frac{dy_i}{d\bar{t}_i}\right)^2 + 3(3 - \varepsilon^3)my_i \\ & - 4\varepsilon^{1+m}\left(\frac{dy_j}{d\bar{t}_j}\right)^2 + 12\varepsilon^{1+m}my_j] / [3(\varepsilon^3 + 1)] \\ & + \delta'_{st}, \end{aligned} \quad (2)$$

where the parameter δ'_{st} denotes random fluctuation. A symmetric ratio of 1.0 (ε =1) makes the first term on the right-hand side of the equation zero, while an asymmetric ratio of $\sqrt[3]{3}$ around 1.5 (ε^3=3) makes the second term zero. The size ratios of 1.00 and approximately 1.50 can be described by the unified number of the n-th root of n.

We define a system as being quasi-stable when only one term on the right-hand side of the differential equation system governing the phenomenon (Eq. (2)) is zero. Life is relatively quasi-stable because d^2y_i/dt^2 becomes smaller when the size ratio of connected parcels takes the values of ε =1 or ε^3 =3. An important clue for clarifying the inevitability and variety of biological molecules is the concept of quasi-stability (meta-stability) weaker than neutral stability. This is because biological systems are essentially neither neutrally stable nor absolutely stable owing to the potential for a brush with death and also since people can live over 50 years while maintaining the mysterious shape of the body. Thus, a new stability concept between the neutrally stable and unstable conditions is necessary for a meta-stable system of life, i.e., quasi-stability. [6-9]

Free nitrogenous bases in water often seem to be a column of two identical bases such as adenine-adenine, which are not in a two-dimensional plane. There are multimeric complexes such as that of guanines. Then, it is well known that identical base pairs of $\varepsilon=1$ are often in RNA, while asymmetric base pairs such as the Watson-Crick type of about $\varepsilon^3=3$ are used in DNA. (Fig. 1)

As Eqs. (1) and (2) show a slightly vague solution for the phenomenon, this indeterminacy also implies that variations of molecular sizes are possible in a limited range. This indeterminacy will permit the possibility of sizes around 1: $\sqrt[3]{3}$, i.e., two types of base pairs such as A-T and G-C.

IV. HIGHER-ORDER OF ANALYSIS

Higher order of the Taylor series

$$f(y)=f(0)+\frac{f^{(1)}(0)}{1!}y+\frac{f^{(2)}(0)}{2!}y^2+\frac{f^{(3)}(0)}{3!}y^3+\ldots\ldots \quad (3)$$

Can be applied to Eq. (1).

The third order of the Taylor series brings

$$\frac{d^2y_1}{d\tau_1^2}=\frac{1}{m_2}\left\{\begin{array}{l} a_{31}y_1+a_{32}y_1^2+a_{33}y_1y_2+a_{34}y_1^2y_2+a_{35}y_1y_2^2+a_{36}y_1^3+a_{37}y_2+a_{38}y_2^2+a_{39}y_2^3 \\ +\left(b_{31}+b_{32}y_1+b_{33}y_2+b_{34}y_1y_2+b_{35}y_1^2+b_{36}y_2^2\right)\left(\frac{dy_1}{d\tau_1}\right)^2 \\ +\left(c_{31}+c_{32}y_1+c_{33}y_1y_2+c_{34}y_1^2+c_{35}y_2+c_{36}y_2^2\right)\left(\frac{dy_2}{d\tau_2}\right)^2 \end{array}\right\} \quad (4)$$

with

$$m_3=9\left(\varepsilon^3+1\right)^3 \square\;\square\; a_{31}=-9\left(\varepsilon^3-3\right)\left(\varepsilon^2+1\right)^2 \square\;\square\; a_{32}=\frac{3}{4}\left(\varepsilon^3-3\right)\left(\varepsilon^6+34\varepsilon^3+33\right)\square\;\square$$

$$a_{33}=12\varepsilon^2\left(3\varepsilon^6-8\varepsilon^4-2\varepsilon^3-8\varepsilon-5\right)\square\; a_{34}=-8\varepsilon^2\left(2\varepsilon^6+7\varepsilon^4+6\varepsilon^3-57\varepsilon-28\right)\square\;\square$$

$$a_{35}=\varepsilon^2\left(-111\varepsilon^6+336\varepsilon^4+426\varepsilon^3+80\varepsilon+25\right)\square\;\square$$

$$a_{36}=\frac{1}{2}\left(\varepsilon^3-3\right)\left(\varepsilon^6-42\varepsilon^3-171\right)\square\; a_{37}=36\varepsilon^2\left(\varepsilon^3+1\right)^2$$

$$a_{38}=-3\varepsilon^2\left(37\varepsilon^6+42\varepsilon^3+5\right)\square\; a_{39}=\varepsilon^2\left(383\varepsilon^6+134\varepsilon^3+7\right)\square\; b_{31}=6\left(\varepsilon^3-1\right)\left(\varepsilon^3+1\right)^2$$

$$b_{32}=-2\left(5\varepsilon^9+5\varepsilon^6-21\varepsilon^3-21\right)\square\;\square\; b_{33}=32\left(\varepsilon^6+\varepsilon^3\right)\square\;\square\; b_{34}=-\frac{32}{3}\varepsilon^3\left(5\varepsilon^3+21\right)$$

$$b_{35}=\frac{8}{3}\left(5\varepsilon^9+9\varepsilon^6-27\varepsilon^3-63\right)\square\;\square\; b_{36}=-\frac{16}{3}\varepsilon^3\left(21\varepsilon^3+5\right)\square\;\square\;\square$$

$$\square\; c_{31}=-12\left(\varepsilon^3+1\right)^2 \square\;\square\; c_{32}=\frac{4}{3}\varepsilon^2\left(-9\varepsilon^6+6\varepsilon^3+15\right)\square$$

$$c_{33}=\frac{32}{3}\varepsilon^2\left(6\varepsilon^6-15\varepsilon^3-5\right)\square\;\square\; c_{34}=\frac{16}{3}\varepsilon^2\left(\varepsilon^6+3\varepsilon^3-14\right)\square\;\square$$

$$c_{35}=32\varepsilon^2\left(2\varepsilon^6+3\varepsilon^3+1\right)\square\;\square\; c_{36}=-\frac{16}{3}\varepsilon^2\left(47\varepsilon^6+41\varepsilon^3+10\right)$$

V. QUASI-STABILITY EXTENDED

We can consider only the disturbance for parcel1, because the system of two parcels is mathematically symmetric. Then, we get

$$\frac{d^2y_1}{d\tau_1^2}=\frac{1}{m_2}\left\{a_{31}y_1+a_{32}y_1^2+a_{36}y_1^3+\left(b_{31}+b_{32}y_1+b_{35}y_1^2\right)\left(\frac{dy_1}{d\tau_1}\right)^2\right\} \quad (5)$$

as the third order of approximation.

The three terms of $a_{31}y_1, a_{32}y_1^2, a_{36}y_1^3$ come from the surface tension, whereas the later terms of $b_{31}, b_{32}y_1, b_{35}y_1^2$ come from convection.

Here, we define

$$a_{31}=0 \; or \; b_{31}=0$$

as the quasistable condition, while

$$a_{32}=0,\; a_{36}=0,\; b_{32}=0,\; or \; b_{35}=0$$

is semi-quasistable.

Table 1 shows the size ratios of parcels, which bring quasistable and semi-quasistable conditions.

The semi-quasistable ratio of about 1: 3.6 will correspond to a threefold variation between 240 of cysteine as the maximum and 75 of glycine as the minimum, while the other semi-quasistable ones of about 1: 1.27 and 1:1.35 lead to some variations of base pairs such as A-T and G-C.

Table 1.Size ratios of parcels (The third order of approximation type)

Stability	surface tension term	convective term
Quasistable	3:2	1:1
Semi-quasistable	3:2	1.27:1
Semi-quasistable	3:2 3.6:1	1.35:1

VII. CONCLUSION

Higher-order of analysis for the cyto-fluid dynamic theory reveals the size ratios of 1:1 and about 2:3 for nitrogenous bases and also the ratios over 2:3 for amino acids..

REFERENCES

[1]Miller S.L, (1953) A production of amino acids under possible primitive earth conditions. Science, 117: 528-529.

[2]Oro J and Kimball A.P, (1961) Synthesis of Purines Under Possible Primitive Earth Conditions. I. Adenine from Hydrogen Cyanide, Arch. Biochem. Biophys. 94: 217–227.

[3]Yanagawa H, (1991) How Life Originated. TBS Britannica. 1-466.

[4]Ricardo A and Szostak J.W, (2009) Origin of Life on Earth. Scientific American.

[5]Watson J.D, Hopkins N.H, Roberts J.W, et al. (1987) Molecular Biology of the Gene, Fourth edition: The Benjamin/Cummings Publishing Company.

[6]Naitoh K, (2001) Cyto-fluid Dynamic Theory, Japan Journal of Industrial and Applied Mathematics, 18-1: 75-105.

[7]Naitoh K, (2003) Macroscopic kinetic model on the origins of biological information. Trans. of JIAM: 115-137.

[8]Naitoh K, (2008) Stochastic Determinism underlying Life, Artificial Life and Robotics, Springer, 13: 10-17.

[9]Naitoh K, (2010) Onto-biology: clarifying the spatiotemporal structure, Artificial Life and Robotics, Springer, 15: 117-127.

[10]Lamb H, (1945) Hydrodynamics, 6th ed. New York: Dover.