
Journal of Biological Physics **28:** 701–712, 2002.

Field Hypothesis on the Self-regulation of Gene Expression

K. YOSHIKAWA
Department of Physics, Graduate School of Science, Kyoto University, Kyoto 606-8502, Japan
(*E-mail: yoshikaw@scphys.kyoto-u.ac.jp*)

Abstract. The mechanism of the self-regulation of gene expression in living cells is generally explained by considering complicated networks of key-lock relationships, and in fact there is a large body of evidence on a huge number of key-lock relationships. However, in the present article we stress that with the network hypothesis alone it is impossible to fully explain the mechanism of self-regulation in life. Recently, it has been established that individual giant DNA molecules, larger than several tens of kilo base pairs, undergo a large discrete transition in their higher-order structure. It has become clear that nonspecific weak interactions with various chemicals, such as polyamines, small salts, ATP and RNA, cause on/off switching in the higher-order structure of DNA. Thus, the field parameters of the cellular environment should play important roles in the mechanism of self-regulation, in addition to networks of key and locks. This conformational transition induced by field parameters may be related to rigid on/off regulation, whereas key-lock relationships may be involved in a more flexible control of gene expression.

Key words: DNA condensation, environmental parameter, first-order phase transition of DNA, higher-order structure of DNA, on/off regulation, segregation in a chain

1. Introduction

One of the most fascinating aspects of living things is their ability to develop autonomously. Even simple prokaryotic cells are capable of self-reproduction, a flexible response and self-management to survive. Modern biology has clarified that life is maintained under genetic control. Higher organisms develop from a single fertilized egg. The result is a highly reproducible spatio-temporal arrangement of differentiated cells, including cell division and differentiation, morphological changes in cells and tissues, locomotion, and apoptosis (programmed cell death). Since the genetic information embedded in DNA is usually the same among all of the cells, a crucial problem is how a living system can exhibit self-regulation using the 'read-only' memory stored as one-dimensional base sequence in DNA.

It is currently considered that the self-regulation of gene expression is performed under complicated networks of biochemical reactions/interactions [1–4]. In biochemical networks, genetic activity is expected to be controlled by molecular signals that determine when and how often a given gene is transcribed, where the

'control parameters' are the intracellular concentrations of chemical species, and the network can be described with these 'control parameters' as coupled ordinary differential equations. Thus, multiplicity in stablle fixed points on the coupled equations correspond to the existence of different states in cells. From mathematical perspective, nonlinearity is necessary to obtain such multi-stability, or a multiple number of fixed points, in coupled kinetic equations.

2. Problems with the Network Hypothesis

The most common concept regarding the mechanism of genetic regulation can be described as follows [1–4]. A protein produced using the information encoded by a specific gene acts as a control factor in regulating the expression of other genes. For example, the concentration of protein A in a cell determines the rate of production, or gene expression, of protein B. The rate of production of protein C is, then given as a function of the concentration of protein B, the rate on protein D is as a function of the concentration of protein C, and so on. Additional signals from other cells can also affect the rates of ongoing biochemical reactions. Since regulatory proteins and other chemical factors may act in combination with other signals to control many other genes, a complicated network with many branches and loops may be generated. In such a network, one regulatory protein can control a gene that produces another regulator, which in turn controls still other genes, where the regulators can be either activators or inhibitors, and the network includes a huge number of positive and negative feedback mechanisms.

Usually, the kinetics of biochemical reactions, including the rate of production of a protein through gene expression, are interpreted within the framework of a mass-action law. If we assume suitable nonlinearity in the differential equation of each kinetic process in the genetic network, we can predict a rich variety of phenomena that are characteristic of a nonlinear dynamical system, such as multiple basins of attraction, temporal rhythm (limit-cycle oscillation), switching (bifurcation), spatiotemporal structures including a Turing pattern, a spiral wave, chaos, etc [5]. Indeed, several studies have addressed such kinetic networks, and many of them have assumed the presence of cubic nonlinearity [6–9] as the product of the concentration of a promoter (or repressor) and the square of the concentration of a regulator. If the network contains a loop and cubic nonlinearity is embedded in the kinetic loop, multiple fixed points, or attractors, are generated. Thus, the switching phenomenon between different states in cells can be interpreted in terms of nonlinear dynamics [10].

Unfortunately, there is a serious problem with this network hypothesis [11]. The framework of the mass-action law may be correct, to the extent that we can examine the reaction kinetics in a test tube with a size on the order of cm $\sim$ mm. In laboratory experiments, the typical concentration of reactants and enzymes is mM $\sim$ nM, where M is the molar concentration. If we consider 1 μM of a reactant or an enzyme in a test tube with a volume of 1 μl, the number of molecules is

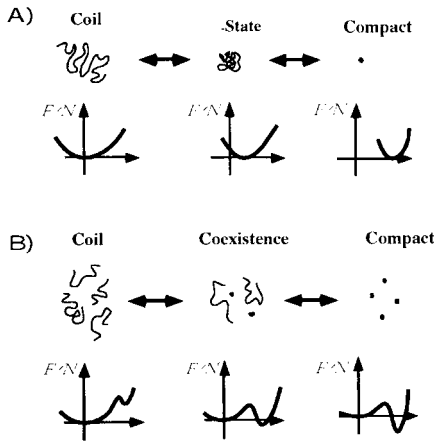

Figure 1. Schematic representation of the conformational transition of giant DNA molecules and the corresponding free-energy profile with respect to the segment density, η. A) Hypothesis most-common until the mid-1990's. B) Correct scheme.

$6{\cdot}10^8$. Thus, in a homogeneous medium, as in usual laboratory experiments, the central limit theorem would almost hold: we can interpret the kinetics by neglecting the stochastic effect, or by ignoring changes in the concentrations of chemicals. In contrast, living cells are typically on the order of μm. In addition, there are thousands of different chemical species in a living cell. For example, human cells contain 30000 genes. It would be unreasonable to expect that there is a sufficient amount of each of the regulatory factors of these 30000 genes in a single human cell on the order of 10 μm to neglect the effect of fluctuation. Actually, several scientists have already noticed the critical effect of such fluctuation on the network hypothesis [12]. Robust on/off switching is difficult to apply under the framework of the network hypothesis.

Considering the present state of the modeling of cellular behavior, Brooks claimed that 'we might be missing something fundamental and unimagined in our models of biology' [11]. In a recent paper, Endy and Brenet stated that 'representation of cellular processes that can be used to compute their future behavior would be general scientific and practical value. But past attempts to construct such representation have been disappointing' [13]. They continued by saying that '... by the early 1970s it was becoming apparent that few if any living systems had complex regulatory networks of this type, and that living systems regulate their transitions from state to state in other ways.' Thus, an important issue for us to consider is 'What are the other ways?'

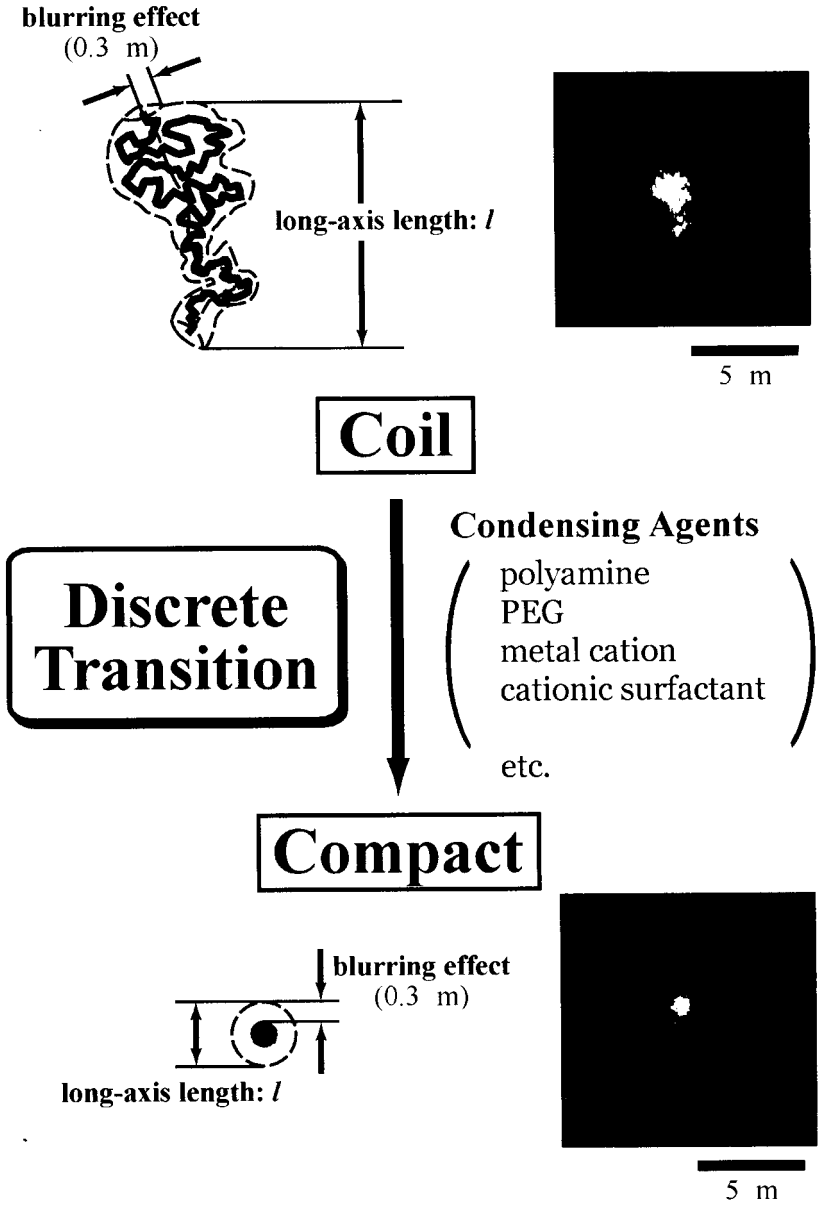

Figure 2. The folding/unfolding transition in giant DNA is largely discrete. Fluorescence images of single T4 DNA, 155 kbp, and the corresponding microscopic structures are shown for both the elongated coil and folded compact states.

3. On/off Switching in Giant DNA

All living cells on Earth posses giant DNA molecules larger than on the order of mega base pairs, M bp. In humans, the total of the full lengths, or contour lengths, of DNA molecules is about 2 m. Even for bacteria, the contour length is on the order of mm. Such long DNA chains are compacted in an intracellular space on the order of μm. Until the mid-1990's, so-called 'DNA condensation' [14, 15] was considered a cooperative phenomenon with a highly cooperative but continuous transition (see e.g., Figure 1A). All of the available experimental data on this transition, such as lightscattering, sedimentation, viscosity, etc., seemed to support the diffuse nature of DNA condensation. Recently, based on the direct measurement of the conformation of single DNA molecules, it has been revealed [16–19] that the transition between an elongated coil state and a folded compact state is largely discrete at the level of individual chains, whereas the physico-chemical properties of an ensemble of chains are always continuous in nature (see e.g., Figure 1 B and Figure 2). The diffuse nature of this transition in the ensemble of molecular chains can be attributed to the presence of a rather wide area for the coexistence of the two states with respect to the concentration of condensing agents. Thus, a giant DNA molecule larger than several tens of kilo base pairs, kbp, is large enough to exhibit an all-or-none behavior, or switching, in the folding transition. It has been confirmed that the discrete nature of the folding transition is rather general and independent of the condensing agent, such as polyamine [17, 19, 20], metal cation (Co(III)[18], Fe(III) [21]), hydrophilic polymer (PEG [16]), surfactant (cationic [22, 23] and neutral [24]), etc.

4. Why Discrete?

As for the discrete nature of the folding/unfolding transition of giant DNA, two factors, stiffness and negative charge, play an important role. The double-stranded structure makes DNA a rather stiff polymer. Double-stranded DNA is about 2 nm in diameter, d, whereas its persistence length is around 50 nm and the Kuhn length, a, is around 100 nm. The ratio of the Kuhn length and the diameter, ξ = a/d, is a measure of polymer stiffness. In double-stranded DNA, $\xi \approx 50$, indicating a high degree of stiffness. Here, we introduce a parameter a as a measure of the degree of polymer swelling, where $\alpha = 1$ corresponds to an ideal chain. For a polymer chain with N (number of Kuhn segments) and α (degree of swelling), the density of the segments, ϕ, is given as [25]

$$\phi \approx \xi^{-1} N^{-1/2} \alpha^{-3} \tag{1}$$

Thus, the number of binary collisions, or the self-avoiding effect, can be written as,

$$\phi \cdot N \approx \xi^{-1} N^{1/2} \alpha^{-3} \tag{2}$$

This relationship indicates that the interaction energy scales as ξ^{-1} $N^{1/2}$ when the segment density is not high, as in the slope in Figure 3A. Recent experiments

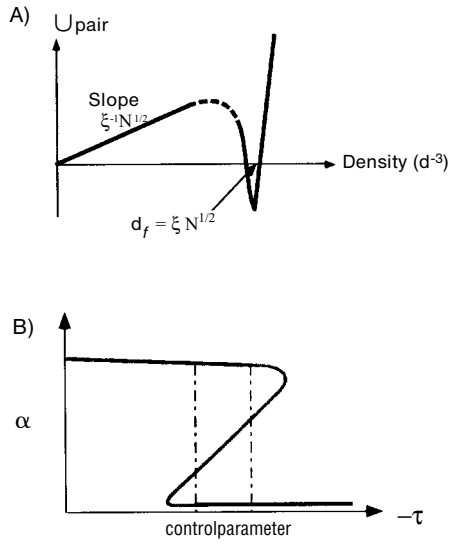

Figure 3. A) Pair interaction energy vs. segment density in a stiff polymer. B) Degree of swelling of a stiff polymer vs. a control parameter.

have confirmed that folded compact DNA exhibits a densely ordered packing of segments, and that the negative charge completely disappears on the inside of the folded product [26]. This means that tightly compact DNA can be considered a kind of ionic crystal. Thus, the swelling ratio in the folded state is given as,

$$\alpha \approx \xi^{-1/3} N^{-1/6} \tag{3}$$

We can depict the profile of the pair interaction U as a function of the density of the segments, α^{-3}, as in Figure 3A. The total free energy, F, of a single polymer chain is given by introducing an elastic term in addition to the above-mentioned interaction term. By analyzing the total free energy, the dependence of the swelling ratio α on reduced temperature τ, or 'field parameter', can be obtained as in the schematic representation in Figure 3B, where α is calculated from the condition $\partial F/\partial \alpha = 0$.

The profile of the change in α fairly well reproduces the general results of experiments on the folding transition of giant DNA molecules. In particular, the size in the coil state remains essentially constant regardless of the change in the field parameter or the concentration of the condensing agent. In actual experiments using single-chain observation, this exact behavior has been observed in DNA molecules, i.e., with a gradual increase in the concentration of the condensing agent, the size of the elongated DNA shows almost no change up to the concentration that causes the folding transition.

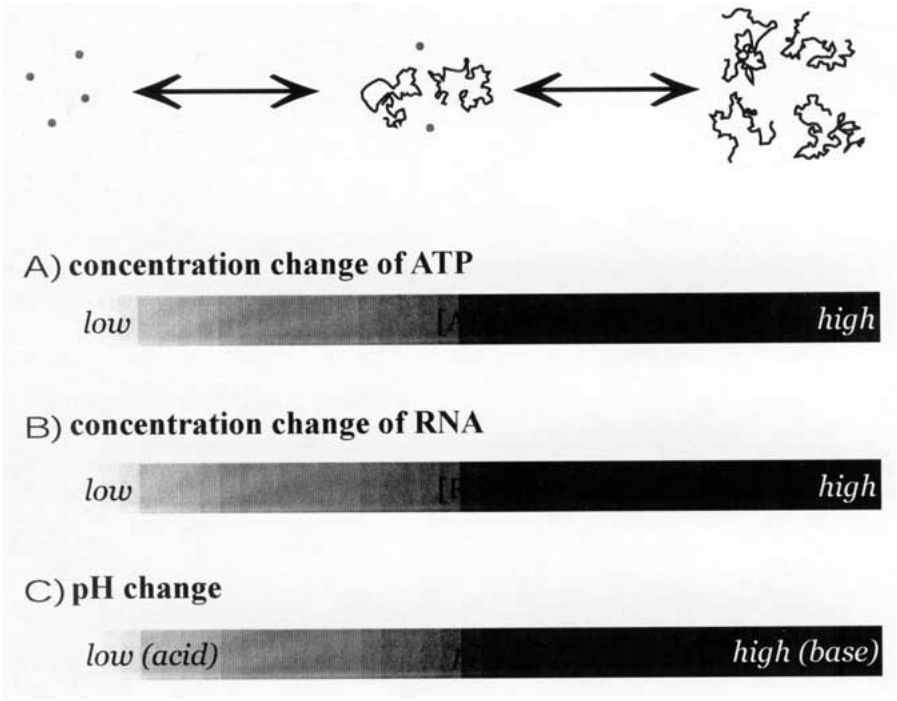

Figure 4. Conformational transition of DNA molecules in the presence of a fixed amount of polyamine. A) An unfolding transition is induced by an increase in the ATP concentration [27], B) the RNA concentration [28], or C) the pH.

5. Cross-Talk Between the Folding Transition of DNA and Environmental Parameters

As noted above, it is evident that giant DNA molecules undergo switching on their conformation. For elongated coil DNAs, aqueous solution is a good solvent because of the high negative charge density along the double-stranded chain. On the other hand, for the folded compact state, almost all of the negative charge on DNA is neutralized due to the association of the phosphate group with a counter cation, which is accompanied by a decrease in free energy due to the decrease in the translational entropy of the counter cation. The large difference in the charge of DNA chains means that a large number of counter ions are absorbed/released with the folding/unfolding transition, respectively [16, 18, 26]. We should note that giant DNA molecules are entrapped within a narrow space in a cell. This implies that cross-talk between the conformational state of DNA and the cytoplasmic environment should have a significant effect in living cells.

Recently, it has been shown from *in vitro* experiments that giant DNA molecules undergo an unfolding transition induced by an increase in ATP in the presence of a fixed amount of spermidine, a natural polyamine (Figure 4) [27]. Polyamines are common chemicals that are found in both prokaryote and eukaryote. A similar

unfolding transition is observed [28] with an increase in RNA in the solution, and also with an increase in pH. None of these chemicals show a specific interaction with DNA. Instead, these species exist in cells in rather high concentrations. We would like to regard the concentrations of these abundant, non-specific chemicals as environmental parameters, as in the parameter τ in Figure 3B. Since these environmental parameters should be involved in the activity of living cells, it is expected that, in a narrow intracellular space, giant DNA molecules should exhibit cross-talk with regard to their respective conformations through these parameters as mediators [29].

6. Dynamic regulation of the Higher-Order Structure of DNA

As has been well established in statistical physics, a one-dimensional Ising model in the absence of long-range interaction can never undergo a phase transition. Giant DNA molecules are one-dimensional chains embedded in three-dimensional space. Thus, giant DNA can undergo a first-order phase transition as explained earlier. In a first-order phase transition, the existence of a phase boundary usually causes a large penalty in the free energy, i.e., the phase-segregated state is unstable. For the ensemble of giant DNA molecules, phase segregation is generated for a rather wide parameter area, since the segregation of different molecular chains does not impose a cost on the interfacial energy. Let us discuss the stability of a segregated state in a single DNA molecule. Since a polymer chain is a one-dimensional string, the boundary between different phases is a zero-dimensional point, if we consider the intrachain coexistence of elongated and compact states Thus, in the phase-segregated state of a polymer chain, the free-energy penalty becomes minimum, in contrast to the usual two- or three-dimensional cases. In addition, the characteristic length of Coulombic interaction in an aqueous environment is on the order of nm, whereas the elastic interaction is effective over the entire chain and the translational entropy of the counter ions is long-range over the solution. In the tightly compact ordered state of DNA, the characteristic distance between segments is on the order of 0.1 nm. Thus, as a result of competition between these long- and short-range interactions and medium-range Coulombic interaction, an intrachain phase-segregated state can have an absolute free-energy minimum within a certain parameter area.

Figure 5 summarizes the nature of the folding transition of a giant DNA molecule induced by various chemical and biochemical compounds. Interestingly, specific, strong binding of a protein such as a restriction enzyme, induces only local disturbance of the conformation [30], whereas nonspecific, weak binding, as with spermidine, causes a discrete transition over the entire molecular chain. The phase segregated state appears between strong and weak binding. Among binding chemicals with intermediate strength, stronger binding with histone H1 [31] causes a partially folded structure on the order of several kilo base pairs while

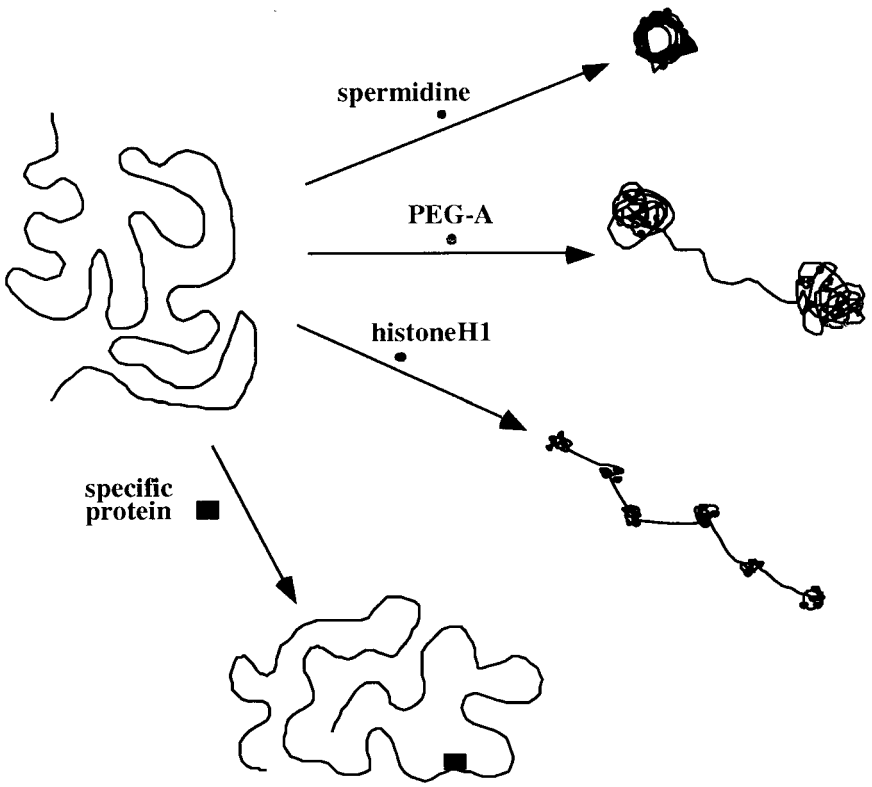

Figure 5. Weak, nonspecific binding induces large-scale changes in the DNA conformation, whereas specifc binding causes only a local deformation in DNA. The characteristic length of the segregation depends on the degree of 'nonspecific interaction'.

weaker binding with PEG-A (amino-pendant polyethylene glycol [32]) causes a larger partially folded structure of several tens of base pairs.

The fact that nonspecific interaction causes global switching in the higher-order structure of DNA likely has biological significance. Currently, the network hypothesis is based on the framework of the 'key-lock' concept. There is abundant evidence that living organisms use the key-lock relationship. However, we would like to stress the importance of non-specific environmental parameters in the mechanism of self-regulation in living things.

7. Higher-Order Structure of DNA vs. Genetic Activity

Previous reports have noted that the genetically active part of chromatin exists in a somewhat relaxed state [33]. This experimental trend has been explained as follows. As promoting factors access a specific domain of DNA, the double-stranded chain is dissolved into a single-stranded state to be transcribed by RNA

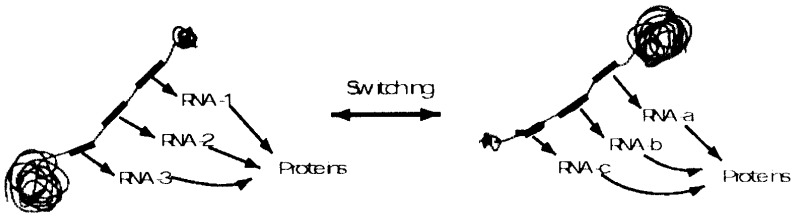

Figure 6. Schematic representation on the switching of gene expression accompanied by the folding/unfolding transition for regions, or holders, of DNA containing multiple genes.

polymerase, and this results in loosening of the packing in chromatin. This explanation does not seem to consider the intrinsic property of DNA to undergo a switching transition. Thus, it may be worthwhile to reconsider the relationship between the structure of chromatin and genetic activity. It is possible that the cause-and-effect relationship regarding the conformation of DNA and genetic activity has been misunderstood. We would like to propose a scenario in which, with a change in an environmental parameter such as the level of ATP, a certain part of chromatin loosens and allows access to transcriptional machinery.

It is also important to indicate the possibility of stepwise unfolding/folding of giant DNA [29]. As has been explained already, the switching transition of giant DNA is inevitably accompanied by the release or absorption of a large number of small ions. For a small system such as in the cytoplasmic space, the transition of DNA molecules is expected to proceed in a stepwise manner, as is schematically shown in Figure 6. Since the scale of the transition in the higher-order structure should be greater than several tens of base pairs, such partial loosening would also be greater than the order of several tens of base pairs. Considering that the sequence of amino acids in a protein corresponds to 100∼300 bases, the partial loosening would involve several tens or several hundreds of genes. This suggests that the conformational transition in a certain part of DNA acts as a 'holder' of genetic information. It is natural to consider such a system that regulates a huge amount of genetic information, i.e., 30,000 genes in the case of humans, as a combination of both individual genes and 'holders' for several genes. Since the key-lock relationship between a specific regulatory chemical and a specific DNA domain is a kind of dynamic equilibrium, such regulation would be flexible, i.e., applicable for continuous control. On the other hand, since the conformational transition of DNA occurs on a larger scale, it would be applicable for switching on a larger number of genes, such as in cell differentiation. Such regulation with a change in the higher-order structure of DNA is based on 're-writable' memory, whereas the one-dimensional base sequence is 'read-only memory'. In relation to

this hypothesis, it is interesting to note a recent finding that highly expressed genes form clusters in chromosomal domains [33].

Further studies are needed to examine the present hypothesis. It should at least be clear that living organisms should not be able to maintain life simply using the 'read-only memory' embedded as a one-dimensional base sequence. The concept of a 'field', as described using environmental parameters, may be useful for, understanding 'what is life?'

References

1. Jacob, F. and Monod, J.: Genetic Regulatory Mechanisms in the Synthesis of Proteins, *J. Mol. Biol.* **3** (1961), 318–356.
2. Britten, R.J. and Davidson, E.H.: Gene Regulation for Higher Cells: A Theory, *Science* **165** (1969), 349–357.
3. Nicolis, G. and Prigogine, I.: *Self-Organization in Nonequilibrium Systems*, John Wiley & Sons, New York, 1977, pp. 354–426.
4. Goldbeter, A.: *Biochemical Oscillations and Cellular Rhythms*, Cambridge University Press, 1996.
5. Murray, J.D.: *Mathematical Biology*, Springer-Verlag, Berlin, 1990.
6. Elowitz, M.B. and Leibler, S.: A Synthetic Oscillatory Network of Transcriptional Regulators, *Nature* **403** (2000), 335–338.
7. Gardner, T.S., Cantor, C.R. and Collins, J.J.: Construction of a Genetic Toggle Switch in Escherichia coli, **403** (2000), 339–342.
8. Hasty, J., Pradines, J., Dolnik, M. and Collins, J.J.: Noise-Based Switches and Amplifiers for Gene Expression, *Proc. Natl. Acad. Sci., USA* **97** (2000), 2075–2080.
9. von Dassow, G., Meir, E., Munro, E.M. and Odell, G.M.: The Segment Polarity Network is a Robust Developmental Module, *Nature* **406** (2000), 188–192.
10. Dziarmaga, J.: Stochastic Gene Expression: Density of Defects Frozen into Permanent Turing Patterns, *Phys. Rev. E* **63** (2000), 011909.
11. Brooks, R.: The Relationship between Matter and Life, *Nature* **409** (2001), 409–411.
12. McAdams, H.H. and Arkin, A.: Stochastic Mechanisms in Gene Expression, *Proc. Natl. Acad. Sci., USA* **94** (1997), 814–819.
13. Endy, D. and Brent, R.: Modeling Cellular Behavior, *Nature* **409** (2001), 391–395.
14. Widom, J. and Baldwin, R.J.: Monomolecular Condensation of DNA Induced by Cobalt Hexamine, *Biopolymers* **22** (1983), 1595–1620.
15. Bloomfield, V.A.: DNA Condensation, *Curr. Opin. Struct. Biol.* **6** (1996), 334–341.
16. Vasilevskaya, V.V., Khokhlov, A.R., Matsuzawa, Y. and Yoshikawa, K.: Collapse of Single DNA in poly(ethylene glycol), Solutions, *J. Chem. Phys.* **102** (1995), 6595–6602.
17. Yoshikawa, K., Takahashi, M., Vasilevskaya, V.V. and Khokhlov, A.R.: Large Discrete Transition in a Single DNA Molecule Appears Continuous in the Ensemble, *Phys. Rev. Lett.* **76** (1996), 3029–3031.
18. Yoshikawa, K., Kidoaki, S., Takahashi, M., Vasilevskaya, V.V. and Khokhlov, A.R.: Marked Discreteness on the Coil-Globule Transition of Single Duplex DNA, *Ber. Bunsen-Ges. Phys. Chem.* **100** (1996), 876–880.
19. Takahashi, M., Yoshikawa, K., Vasilevskaya, V.V. and Khokhlov, A.R.: Discrete Coil-Globule Transition of Single Duplex DNAs induced by Polyamine, *J. Phys. Chem. B* **101** (1997), 9396–9401.
20. Yoshikawa, Y. and Yoshikawa, K.: Diaminoalkanes with an Odd Number of Carbon Atoms induce Compaction of a Single Double-Stranded DNA chain, *FEBS Lett.* **361** (1995), 277–281.

21. Yamasaki, Y. and Yoshikawa, K.: Higher Order Structure of DNA Controlled by the Redox State of Fe^{2+}/Fe^{3+}, *J. Am. Chem. Soc.* **119** (1997), 10573–10578.
22. Mel'nikov, S.M., Sergeyev, V.G. and Yoshikawa, K.: Discrete Coil-Globule Transition of Large DNA by Cationic Surfactant, *J. Am. Chem. Soc.* **117** (1995), 2401–2408.
23. Mel'nikov, S.M., Sergeyev, V.G., Yoshikawa, K., Takahashi, H. and Hatta, I.: Cooperativity or Phase Transition? Unfolding Transition of DNA Cationic Surfactant Complex, *J. Chem. Phys.* **107** (1997), 6917–6924.
24. Mel'nikov, S.M. and Yoshikawa, K.: First-Order Phase Transition in Large Single Duplex DNA induced by a Nonionic Surfactant, *Biochem. Biophys. Res. Commun.* **230** (1997), 514–517.
25. Grosberg, A.Yu. and Khokhlov, A.R.: *Statistical Physics of Macromolecules*, AIP Press, N.Y., 1994.
26. Yamasaki, Y., Teramoto, Y. and Yoshikawa, K.: Disappearance of the Negative Charge in Giant DNA with a Folding Transition, *Biophys. J.* **80** (2001), 2823–2832.
27. Makita, N. and Yoshikawa, K.: ATP/ADP Switches the Higher-Order Structure of DNA in the Presence of Spermidine, *FEBS Lett.* **460** (1999), 333–337.
28. Tsumoto, T. and Yoshikawa, K.: RNA Switches the Higher-Order Structure of DNA, *Biophys. Chem.* **82** (1999), 1–8.
29. Takagi, S. and Yoshikawa, K.: Stepwise Collapse of Polyelectrolyte Chains Entrapped in a Finite Space as Predicted by Theoretical Considerations, *Langmuir* **15** (1999), 4143–4146.
30. Oana, H., Ueda, M. and Washizu, M.: Visualization of Specific Sequence on a Single Large DNA Molecule using Fluorescence Microscopy based on a New DNA-Stretching Method, *Biochem. Biophys. Res. Commun.* **265** (1999), 140–143.
31. Yoshikawa, Y., Velichiko, Yu.S., Ichiba, Y. and Yoshikawa, K.: Self-Assembled Pearling Structure of Long Duplex DNA with Histone H1, *Eur. J. Biochem.* **268** (2001), 2593–2599.
32. Yoshikawa, K., Yoshikawa, Y., Koyama, Y. and Kanbe; T.: Highly Effective Compaction of Long Duplex DNA Induced by Polyethylene Glycol with Pendant Amino Groups, *J. Am. Chem. Soc.* **119** (1997), 6473–6477.
33. Wolffe, A.: *Chromatin, Structure & Function*, 2nd ed., Academic Press, 1995.