Molecular computing : Steps toward integration

(分子コンピューティング：実現へのステップ)

Michael Conrad and Klaus-Peter Zauner

Oyo Buturi (Journal of The Japan Society of Applied Physics)
Volume 64 Number 10, pp. 1002--1006, 1995.

応用物理　第64巻　第10号（1995）抜刷

■ 解　説

Molecular computing : Steps toward integration

(分子コンピューティング：実現へのステップ)

Michael Conrad and Klaus-Peter Zauner

著者らは，たんぱく質や核酸などのマクロな分子が持っている形状認識能力に注目し，「統計的集団ではなく個々の分子が重要な機能的役割を果たすような情報処理システム」（分子コンピューター）の研究に取り組んできた．したがって，ここで述べているコンピューティング（計算）は，集積回路などを用いたフォンノイマン型計算ではなく，個々の分子の機能や相互作用を利用した非フォンノイマン型計算を指している．現在の段階では，計算は，さし当たり一般的な計算能力を実証するものではなく，個々の具体的な問題を有効に解くことができることを例証する段階にあるだろう．

さて，分子のレベルのミクロな現象を計算に利用するためには，これをマクロな世界と結びつけることが必要となるが，著者らはこれを変換と増幅（Transduction and Amplification）の並列的な連鎖を用いて行うことを提案している．また，非常に速い分子の形状認識過程には，分子の異なる状態が並行して存在する量子力学的な過程が有効に作用しているとも述べている．

分子コンピューティングの可能性を示唆する具体例として，以下の四つの代表的なプロトタイプシステムを用いて，異なったモードでの計算が可能なことを示している．

1．脂質膜を用いたバイオセンサー．このセンサーの動作には，変換と増幅の原理が使われている．光敏感な媒質との組み合わせを提案している．

2．バクテリオロードプシンを用いた光計算．光メモリー，ホログラフィックプロセシングなど．

3．DNA 鎖を用いた計算．DNA 鎖の連結反応などのバイオテクノロジーを用いてハミルトンパス問題などを解くことが可能なことを示した Adleman の研究を紹介している．本手法に基づく一般的な計算の可能性を指摘している．

4．神経を変換と増幅の構成単位とした仮想的な脳のシミュレーション．

最後に，要素技術・素子の開発とコンピューティングの質の向上が相乗的に進むという分子コンピューティングのシナージェティックな研究の進展を展望している．　（論文紹介　岩崎　裕）

Keywords : **molecular computing, molecular recognition, self-assembly, micromacro interface, biosensors, optical computing, bacteriorhodopsin, DNA computing, neuromolecular computing, biological information processing**

1. Introduction

Molecular computers are information processing systems in which individual molecules, as opposed to statistical aggregates, play a critical functional role[1]. On this definition biological systems are natural molecular computers, and indeed the study of biological information processing and control capabilities from this point of view is one direction of molecular computing research. Another direction is technological. The two directions are highly synergistic ; analysis helps with synthesis and synthesis provides a powerful means of analysis.

The field of molecular computing reaches beyond biologically motivated approaches, since it allows for the development of entirely novel systems that utilize molecular materials. Biotechnology in particular affords the possibility of producing a vast range of new molecular materials that could contribute to the creation of information processing devices, including sensors, measuring devices, and actuators, with performance characteristics that are unattainable by conventional materials.

It is convenient to divide present day molecular computing research into three main directions : shape based, conventional computer mimicry, and optomolecular. Our brief review of recent developments in the molecular computing field will emphasize approaches based on biology-like molecules.

2. Blomolecular essentials

The functionality of biological systems depends on

Department of Computer Science
(5143 Cass Ave, State Hall Rm 431 Detroit MI 48202, USA)

分類番号　12.3，9.3

macromolecular structures that support highly specific functions. The number of possible structures is enormous. However, in all cases they are assembled from a small set of common building blocks. Proteins provide the most striking example. In nature these are built up from twenty types of amino acids, typically linked together in chains 200-400 amino acids in length that fold up to form a characteristic three dimensional shape (in the size range of 10 nm). The folded shape largely determines the ability of the protein to recognize other molecular structures. The image of a key fitting into a lock is often used to characterize this shape-based mode of recognition. Recognition can lead to catalytic switching or to the formation of stable polymacromolecular structures (principle of self-assembly). In both cases the interacting molecules form a supermolecular complex, the difference being that in the case of enzymes the complex must destabilize rapidly in order to combine high specificity with high catalytic turnover rate.

Nucleic acids, like proteins, are built up from a small set of building blocks, in this case the nucleotide bases. For the present purposes the important point is that complementary nucleic acid strands can line up (or anneal) in the correct manner with remarkable rapidity. The process is cooperative, with a phase transition between the self-assembled and the disassembled state.

Proteins are building blocks for higher order structures, such as the cytoskeleton (an active network of fibers important for the structure and movement of the cell). They are also integrated into lipid bilayers which form the membranes that envelop the cell and extend through it. The functional capabilities of proteins can be modulated through other macromolecules and the local physio-chemical milieu. In the laboratory it is possible to enlarge the set of building blocks, both at the amino acid and at the macromolecular level, and to organize them into structures that are not found in natural systems.

3. Transduction-amplification principle

In order to utilize molecular level mechanisms for information processing it is necessary to connect them to macroscopic input and output. Biological systems solve this problem through a variety of ubiquitous transduction-amplification cascades that serve as interfaces across different levels of scale. The principle is

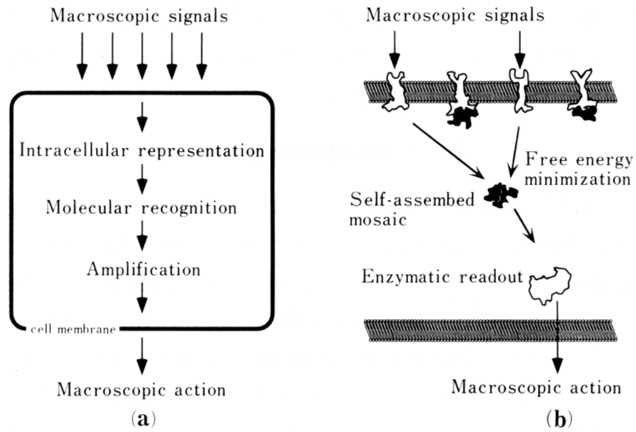

Fig. 1 Schematic illustration of a transduction-amplification module(a) and realization through self-assembly mechanism(b)).

schematically illustrated in **Fig. 1 (a)**. Macroscopic signals impinging on the external membrane of a cell are transduced to internal molecular representations that can be recognized and specifically acted on by macromolecules. The actions eventually culminate in a macroscopic response of the cell. Thus hormones impinging on the cell commonly influence the pattern of gene activation. As another example, nerve impulses impinging on a neuron can trigger the production of second messenger molecules, leading to a sequence of molecular processes that ultimately control the firing of the neuron[2].

The role of shape-based pattern recognition in the transduction-amplification process is illustrated by the conceptual self-assembly device[3] in Fig. 1 (**b**). External macroscopic signals are re-represented as molecular shapes. These self-assemble to form a polymacromolecular mosaic on the basis of free-energy minimization. Different patterns of input signals trigger the appearance of different combinations of molecules, yielding mosaics with shape features common to different groups of input patterns. Readout enzymes then link these shape features to appropriate actions of the cell. In this way a symbolic pattern recognition problem is converted to a free energy minimization process. The cell essentially "crystallizes" the solution to the pattern recognition problem.

In biological cells the principle is embedded in complex webs of molecular processes that combine conformational changes with chemical reactions. The higher level structural framework of the cell —— the membrane and cytoskeletal organization —— probably also plays a signal integrating role[4].

4. Quantum speedup principle

How fast are shape recognition processes of the type that occur in the self-assembly model? Enzymatic processes typically occur on .1 to 1 ms time scale. The time required for self-assembly complex formation is more variable, but it can be as fast as enzymatic recognition. Such times may appear long in comparision to semiconductor switching. However, the number of digital switching operations that would be required to duplicate the molecular pattern recognition capability would be astronomical.

What is the basis of this unintuitive combination of specificity and speed? Macromolecules are large enough so that weak (e. g., van der Waal's) interactions allow for binding forces to become comparable to thermal energies only when a close fit occurs and small enough to explore each other's shapes through diffusional (or Brownian) search. However, it is necessary to add an active principle, consistent with the fact that proteins dynamically change shape during the recognition process. Such nonthermal motions allow the protein to draw in a complex partner and to actively release it. The free energy surface of the reaction must be dynamic, roughly analogous to the manner in which the potential energy of an elastic collision changes dynamically. It means that the complex formation process should be governed by a nonlinear dynamics allowing instabilities. It is likely that the interaction between loosely bound (non-Born-Oppenheimer) electrons and the nuclear coordinates —— the so-called electronic-conformational interaction —— plays an important role. The idea is that the nuclei are agitated by interference among electronic states of slightly different energy and this agitation in turn exerts a perturbing effect on the electronic states. In this way the superposition of electronic states serves to funnel thermal energy into selected degrees of freedom of the nuclear coordinates[5].

The inherent parallelism of the electronic wave function is converted to speed of molecular pattern recognition. This in turn, allows for enhanced pattern recognition capabilities at the cell (or device) level. In effect the entirely picturable macroscopic perception-action capabilities draw power from the nonpicturability of microphysical processes at the submolecular level.

With simple pattern processors, e.g., NAND gates, it is possible to perform any computational function. A fortiori it is possible to build any computational function from more powerful recognizers, such as those that use the transduction-amplification principle. The implementations can be far more efficient, however. Such systems are not programmable in the conventional sense. Learning procedures like evolution are required to mold the dynamics and the interconnections for the desired functions.

5. Prototypes

Let us now briefly review four prototype systems that capture aspects of biomolecular processes and in particular let us consider the character of the computations performed.

5.1 Biosensors

The principle of the biosensor is essentially the same as the transduction-amplification principle. The biosensing device is composed of a molecular recognition unit and a transducer that converts a chemical signal into an optical or electrical output. Possibilities include fluorescent markers, voltage sensitive dyes, charge transfer from proteins to the gate electrode of a field effect transistor and direct contact to proteins through conducting polymers[6]. Bilayer membranes allow for immobilizing the molecular recognition unit without interfering with its function[7]. The important feature of biosensors is that they can react selectively to the presence of specific molecules in a complex medium. The recognition of molecule combinations should also be possible in active media where signal carrying molecules interact. Light sensitive active media have been used in optical pattern processing, though not yet in conjunction with biosensors[8].

The self-assembly device illustrated in Fig. 1 can be thought of as a plausible future elaboration of the same technology that is being developed in conjunction with biosensors. Devices operating on this principle could efficiently process signal patterns that are variable but too context sensitive to be broken up into individually recognizable parts. This follows from the feature that partially formed complexes serve as processors for missing parts, and do so regardless of the time order of their formation. The self-organizing dynamics that confers this extra power precludes conventional programming, but is well suited for learning through variation-selection methods or error feedback signals that act on the internal structure.

5.2 Bacteriorhodopsin-based optical computing

Interfacing macroscopic with microscopic (i. e., transduction and amplification) has been approached in a rather different way using the exceptionally stable protein bacteriorhodopsin (BR). In this light driven proton pump all processes required for signal transduction, amplification, and resetting are built into a single molecule. On light absorption BR switches among eight states with characteristic absorptions and lifetimes (ps to ms). These features, which have been enhanced through genetic engineering, allow for optical memories[9,10] and real time holographic processing[11]. Other applications (sensing, neural network designs) are based on a rapid (picosecond time scale) charge displacement that accompanies one of the state transitions[12].

5.3 DNA string processing

In the 1970's, Vaintsvaig and Liberman proposed a universal stochastic computation scheme based on the ability of enzymes to recognize and act on specific subsequences of DNA[13]. Recently Adleman has utilized the self-assembly of complementary DNA strands to solve a search problem[14]. The problem might be to find a path that visits every vertex in a directed graph given a start and end point (see **Fig. 2**). The vertices and edges of the graph are encoded in a collection of short DNA sequences (this corresponds to the transduction step). The collection is then allowed to interact, leading to the formation of DNA hybrids for each possible pathway. One of the hybrids should encode the solution (if it exists). The solution is amplified through selective polymerase chain reactions (PCR), gel electrophoreses, and affinity-purifications. The transduction and amplification steps are inefficient compared to conventional computing. However, the system affords an enormous amount of fine grained parallelism, connected with the extraordinary affinity of complementary DNA strands, and consequently the approach may be useful for large, highly parallelizable problem domains.

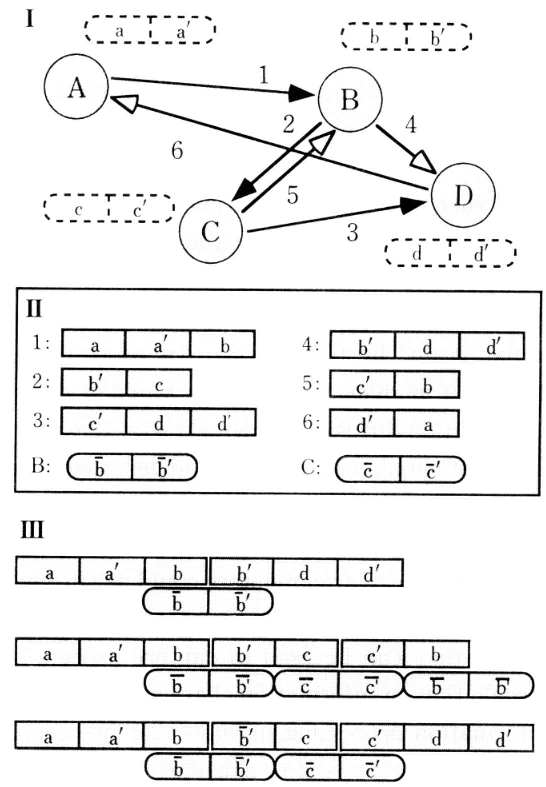

Fig. 2 DNA string processor. The problem (I) is to find a Hamiltonian path (indicated by solid arrows). Each vertex is formally assigned a short and unique DNA sequence (in dashed boxes). A DNA sequence is derived for each edge from this formal assignment (II). The first half of each edge sequence corresponds to the second half of the sequence (indicated by primes) formally assigned to the vertex from which it originates. (See edge 2 and note that the start A and end D are exceptions.) The second half of each edge sequence corresponds to the first half of the sequence formally assigned to the vertex which it enters. The edge sequences and the complements (indicated by bars) of the sequences assigned to the intermediate vertices are synthesized. This set of sequences is allowed to interact. The possible pathways are represented by the reaction products (III), only three of which are shown. If a Hamiltonian path exists the corresponding DNA sequence can be extracted by standard biochemical procedures.

For problems that grow exponentially the amount of DNA that the Adleman system would have to use would also grow exponentially. The nice feature of the system is that it draws on the clever physics of self-assembly to implement efficient pattern matching. As a programmable system it freezes out many of the interactions that biological systems use for problem solving and that could conceivably be used by technological systems.

5.4 Neuromolecular simulations

Finally we can briefly note that our group has used conventional computers to simulate molecular computer designs[15]. The basic idea is to build a virtual "neuromolecular brain" by treating neurons as transduction-amplification modules with various kinds of dynamics, such as reaction-diffusion or cytoskeletal dynamics. A repertoire of simulated dynamic modules is evolved and knit together in a higher level network architecture suitable for performing perception-action tasks. Such virtual neuromolecular computers could never duplicate the recognition capabilities of biological systems in real time. But they can be used to capture aspects of the cross-scale flow of information characteristic of biological information processing in ways that are useful for practical applications and to elucidate the architectural features that would most effectively utilize the recognition capabilities of biomolecular components.

6. Towards integrative synergy

The early 1970's saw the formulation of a variety of molecular computing concepts. The 1980's saw an increasing interest in the possibilities of this field, and increasing efforts to develop workable designs on paper. The possibility of actual prototypes seemed far off. Today there is a spectrum of prototypes, with BR even having reached the marketplace. Each of these prototypes —— only a few of which have been noted here —— captures some fragment of the unique capabilities of carbon. Biosensors utilize molecular specificity, BR utilizes molecular state change, the Adleman system employs the self-assembly properties of naked DNA, and the virtual molecular computer designs capture the vertical flow of information and the learning algorithms that are capable of exploiting this vertical flow. With these prototypes we can begin to see how different biomolecular materials and processes contribute to the capabilities of biological systems, and we can begin to see how in the coming period they can be knit together to yield technological systems that enjoy some of these remarkable capabilities.

Acknowledgement

This material is based upon work supported by the National Science Foundation under Grant No. ECS-9409780.

References

1) M. Conrad: *Advances in Computers*, ed. M. C. Yovits, p. 235 (Academic Press Inc., Boston, 1990).
2) E. A. Liberman, S. V. Minina, O. L. Mjakotina, N. E. Shklovsky-Kordy and M. Conrad: Brain Res. **338**, 33 (1985)
3) M. Conrad: Computer IEEE **25**, 11 (1992).
4) G. Matsumoto, M. Ichikawa, A. Tasaki, H. Murofushi and H. Sakai: J. Membrane Biol. **77**, 77 (1984).
5) M. Conrad: Chaos, Solitons & Fractals **4**, 423 (1994).
6) M. Aizawa: *Introduction to Molecular Electronics*, ed. M. C. Petty, M. R. Bryce and D. Bloor, p. 295 (Edward Arnold, London, 1995).
7) A. Ottova-Leitmannova, T. Martynski, A. Wardak and H. T. Tien: *Molecular and Biomolecular Electronics*, ed. R. R. Birge p. 439 (American Chemical Society, Washington, 1994).
8) N. G. Rambidi: BioSystems **31**, 3 (1993).
9) N. N. Vselodov and G. R. Ivanitskii: Biophysics **30**, 962 (1985).
10) R. R. Birge: Computer IEEE **25**, 56 (1992).
11) N. Hampp, R. Thoma, D. Oesterhelt, and C. Bräuchle: Applied Optics **31**, 1834 (1992).
12) F. T. Hong: BioSystems **19**, 223 (1986).
13) M. N. Vaintsvaig and E. A. Liberman: Biofizika **18**, 939 (1973).
14) L. M. Adleman: Science **266**, 1021 (1994).
15) J.-C. Chen and M. Conrad: Physica D **75**, 417 (1994).

(Received May 24, 1995)