International Journal of Computer Applications Technology and Research
Volume 4– Issue 3, 175 - 178, 2015, ISSN:- 2319–8656

A Study on DNA based Computation and Memory Devices

Harmeet Singh
Department of Computer Science
Punjab College of Technical Education (PCTE)
Baddowal, Punjab, India

Sangeeta
Punjab Technical University, Jalandhar

Harpreet Kaur
Department of Electronics and Communication Engineering
Sant Longowal Institute of Engineering and Technology (SLIET) Longowal,
Punjab, India

Abstract: The present study delineates Deoxyribonucleic Acid (DNA) based computing and storage devices which have good future in the vast era of information technology. The traditional devices mostly used are made up of silicon. The devices are costly and have physical limitations to cause leakage of electrons and circuit to shorten. So, there is a need of materials which are capable of doing fast processing and have vast memory storage. DNA which is a bio-molecule has all these characteristics capable of providing ample storage. In classical computing devices, electronic logic gates are elements which allow storing and transforming of information. Designing of an appropriate sequence or a net of “store” and “transform” operations (in a sense of building a device or writing a program) is equivalent to preparing some computations. In DNA based computation, the situation is analogous. The main difference is the type of computing devices since in this new method of computing instead of electronic gates, DNA molecules have been deployed for the processing of dossier. Moreover, the inherent massive parallelism of DNA computing may lead to methods solving some intractable computational problems. The aim of this research study is to analyze the logical features and memory formation using DNA bio molecules in order to achieve proliferated speed, accuracy and vast storage.

Keywords: DNA; information technology; nanotechnology; bio- molecules; DNA computing

1. INTRODUCTION

Computing is ordinarily defined as the use of computer hardware and software to ameliorate the speed and accuracy of mathematical calculations and manipulations. It is the computer-specific part of information technology. Moore's law portrays a long-term trend in the history of computing hardware, in which the number of transistors that can be placed inexpensively on an integrated circuit has doubled approximately every two years. The silicon chip which has supplied several decades’ worth of remarkable features proliferates in computing power and speed. As silicon computer circuitry gets even smaller in the quest to pack more components into smaller areas on a chip, ultimately the miniaturized electronic devices are undermined by fundamental physical limits. They start to become leaky, making them incapable of holding onto digital information. Many researchers are working to overcome the dilemma of silicon chip technology [1]. In order to overcome the flaws of the current scenario, there is a need of better performance material to process and store the information. DNA bio molecules which are genetic materials have the capability to store and process large amount of data. The concept of computing using DNA was initialized by Leonard Alderman who solved the Hamiltonian path problem using DNA molecules [2]. Nowadays, the research in the area of DNA computing has been continued in designing algorithms, designing new basic operations, developing new ways of encoding information in DNA bio molecules and reduction of errors in computations based upon DNA [3]. This paper affianced is to study and discuss the concept of DNA, logical operations based upon DNA and data storage, its main advantages delineating the crucial role of DNA computing in the field of information technology.

2. FOUNDATION OF DNA COMPUTING

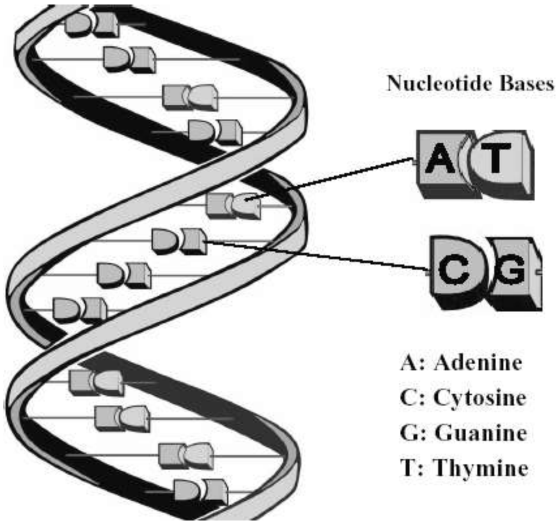

Figure.1 The structure of DNA double helix

2.1 Concept of DNA

Deoxyribonucleic acid (DNA) is a nucleic acid that incorporates the genetic instructions used in the development and functioning of all known living organisms and some viruses. The fundamental role of DNA molecules is the long-term storage of information. DNA is oftentimes compared to a set of blueprints or a recipe, or a code, since it contains the instructions required to construct other components of cells, such as proteins and RNA molecules. The DNA segments that carry this genetic information are called genes, but other DNA sequences have structural purposes, or are involved in regulating the use of this genetic information.

Figure 1 shows DNA double helix structure which is a double stranded sequence of four nucleotides; the four nucleotides that compose a strand of DNA are as follows: adenine (A), guanine (G), cytosine (C), and thymine (T); they are intermittently called as bases. The chemical structure of DNA (the famous double- helix) was discovered by James Watson and Francis Crick in 1953. It consists of a particular bond of two linear sequences of bases. This bond follows a property of complementarity: adenine bonds with thymine (A-T) and vice versa (T-A), cytosine bonds with guanine (CG) and vice versa (G-C). This is known as Watson-Crick complementarity.

2.2 General Working of DNA Computation

DNA is the dominant information storage molecule in living cells. On behalf of using electrical impulses to represent bits of information, the DNA computer adopts the chemical properties of these molecules by examining the patterns of combination or growth of the molecules or strings. DNA can do this through the manufacture of enzymes, which are biological catalysts that could be called the 'software', used to accomplish the desired calculation. DNA computers use deoxyribonucleic acids A (adenine), C (cytosine), G (guanine) and T (thymine) as the memory units and recombinant DNA techniques already in existence carry out the fundamental operations. From computer science point of view, a DNA strand is a word over alphabet ΣDNA = {A, C, G, T}. In a DNA computer, computation takes place in test tubes or on a glass slide coated in 24K gold. The input and output are both strands of DNA, whose genetic sequences encode certain information. A program on a DNA computer is executed as a series of biochemical operations, which have the effect of synthesizing, extracting, modifying and cloning the DNA strands.

2.3 Basic Operations

Hence, the basic operations of DNA algorithms are usually constructed for selecting sequences which satisfy some particular conditions. On the other hand, there may be different sets of such basic operations. In fact, any biochemical procedure which may be interpreted as a transformation (or storing) information encoded in DNA molecules may be treated as a basic operation of DNA based algorithms. One of the possible set of such operations is the following [4]:

MERGE: given two test tubes N1 and N2 create a new tube N containing all strands from N1 and N2.

AMPLIFY: given tube N create a copy of them.

DETECT: given tube N return true if N contains at least one DNA strand, otherwise return false.

SEPARATE: given tube N and word w over alphabet ΣDNA create two tubes +(N, w) and !(N, w), where +(N, w) consists of all strands from N containing w as a substring and: (N, w) consists of the remaining strands.

LENGTH-SEPARATE: given tube N and positive integer n create tube (N, ≤ n) containing all strands from N which are of length n or less.

POSITION-SEPARATE: given tube N and word w over alphabet ΣDNA create tube B(N, w) containing all strands from N which have w as a prefix and tube E(N, w) containing all strands from N which have w as a suffix. Each of the above operations is a result of some standard biochemical procedure.

3. MOLECULAR INFORMATION STORAGE

As in magnetic information storage, where magnetic states of ferromagnetic compounds are used, electrochemical information storage is also being studied in biological systems. In this case, distinct oxidation states of certain complex chemicals are being used for multiple bits of information storage at the molecular level. In principle, the amount of information stored is directly related to the number of oxidation states obtainable. By converting the redox state of the molecules into electrical signals, the information can be easily read. This is accomplished by allowing the chemical complexes to self-assemble into monolayers on gold electrodes. This technology is an example of how, as used in Dimensional Design, different kinds of input and output stimuli (chemical input and electrical energy as the output) can be used to store and read information.

A variety of complex “triple-decker” complexes of ferrocenes and porphyrins have been recently synthesized which have four stable and distinct redox states [5]. This is another example of how to build systems which can store enormous amounts of information.

The genetic code is a prime example of how biological systems store an enormous amount of information at the molecular level. Biological information storage in DNA is based on a genetic alphabet. The information content is believed to be encoded in two base pairs (A and G or C and G) which hold the two strands together. In fact there is a special type of bond (the hydrogen bond) between the two base pairs which holds the DNA molecule together. Genetic information is encoded in the specific sequences, along the DNA, of the A-C base pair and the C-G base pair. The number of bits of information is directly related to the sequences of base pairs. The genetic information is then “read” by a complex series of interactions between DNA and special enzymes and proteins. The output of the system is the generation of a replicate strand of DNA (as in cell replication) or the generation of specific amino acids, which are then assembled into specific proteins. Deciphering of the information stored in the base sequences of the genetic code allows biological systems to survive. Although DNA uses a four letter alphabet, DNA is fundamentally a binary storage media for imprinting and retrieving chemical information.

Although DNA has a binary storage capacity, the information content in DNA is so vast we can only conclude that another mechanism, in addition to the known and previously described information storage mechanisms must be at work here. One example of the enormity of the information in DNA is the fact that DNA contains output information not only about specific amino acids and bases, but also about how these building blocks are organized into complex three dimensional (3D) bio-molecular structures. In order to understand and utilize the enormous information in DNA the new field of bio-informatics has developed which uses complex mathematical modeling, algorithm-based computer technology, DNA chips and cDNA microarrays (DNA coated microcircuits) [6]. A high-throughput sequencing ability is characteristic of this new sophisticated genomic computational technology.

4. DNA BASED LOGIC GATES

To build a computational system, it is firstly necessary to have the development of DNA based logic devices. A logic gate performs a logical operation on one or more logic inputs and produces a single logic output. The logic normally performed is Boolean logic and is most commonly found in digital circuits. Figure 2 shows the block diagram of a simple logic device with two inputs and one output.

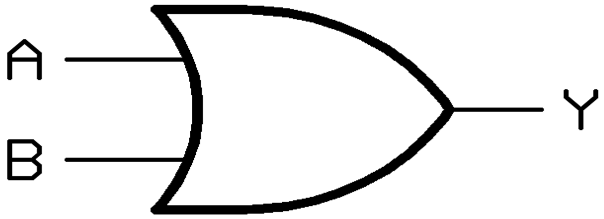

Figure.2 Schematic of Logic gate with input and outputs

In vitro studies have been used to design combinations of molecules that have emergent properties related to information processing--molecular computing devices. Both the inputs and outputs consist of molecular species, with the output being a biologically active molecule. The extent to which these devices will be used with the cellular context is unclear--however, they are bound to inspire new directions for research in synthetic biology, and have potential applications in biochemical sensing, pathway engineering, and medical diagnosis and treatment.

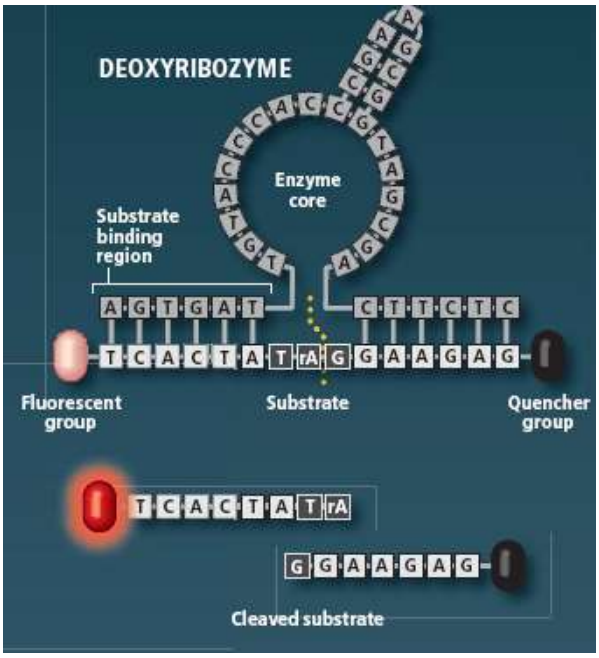

Figure.3 Schematic of DNA enzymes combination

Combining DNA enzymes with stem-loop controllers yields a variety of fundamental logic gates that use short strands of DNA as both inputs and outputs. The cleaving action of the enzyme produces the strands that serve as the gate's output of 1 as shown in Figure 3.

4.1 AND Gate

A logical AND gate has two inputs and produces an output of 1 only if both inputs are 1. A deoxyribozyme with a stem-loop on each of its arms acts as an AND gate. Figure 4 shows the working of AND gate made with DNA [7] .The closed stems disable the enzyme (left), and only when both loops' matching input strands are added can the enzyme cleave substrates (middle). Truth table (right) summarizes the gate's function.

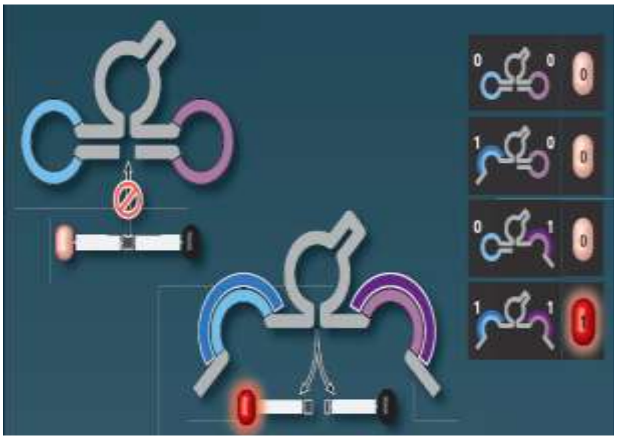

Figure.4 Working of DNA based AND gate

4.2 AND-AND-NOT Gate

A stem-loop controller on the "back" of a deoxyribozyme acts as a NOT input that inhibits the enzyme when the matching input strand is present. If the stem-loop's input strand is not present (0), the stem remains closed and the enzyme cleaves substrates to produce output strands, provided that the enzyme's arms are free (left). Figure 5 shows working of AND-AND-NOT Gate. When the input strand binds to the controller, the stem opens, deforming the enzyme core and rendering it inactive (middle). A deoxyribozyme with controllers on both arms and its back thus behaves as an AND-AND-NOT gate. The enzyme is active, cleaving substrates and thus producing the 1 output, only if inputs X (blue) AND Y (purple) AND NOT Z (yellow) are present.

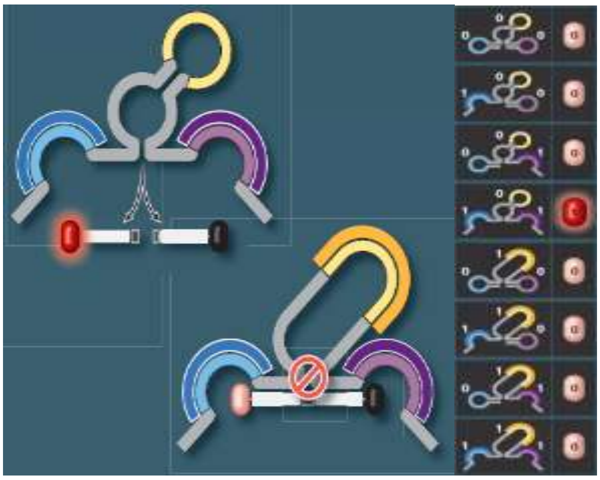

Figure.5 Working of DNA based AND-AND-NOT gate

5. DNA'S INFORMATION STORAGE AND PROCESSING CAPABILITIES

Nucleic Acids are used because of density, efficiency and speed. DNA molecules can store far more information than any existing computer memory chip. It has been estimated that a gram of DNA can hold as much information as a trillion CDs.

Most electronic computers operate linearly and they manipulate one block of data after another, biochemical reactions are highly in parallel: a single step of biochemical operations can be set up so that it affects trillions of DNA strands. While a DNA computer takes much longer than a normal computer to perform each individual calculation, it performs an enormous number of operations at a time and requires less energy and space than normal computers.

6. CONCLUSIONS

The field of DNA computing remains alive and promising, even as new challenges emerge. Most important among these are the uncertainty, because of the DNA chemistry, in the computational results, and the exponential increase in number of DNA molecules necessary to solve problems of interesting size. So, with the commencement of this field, lots more are expected in the betterment of Information technology. In addition, new paradigms based on molecular evolution have emerged from molecular biology to inspire new directions in DNA computing. As has been the case in the recent development of new fields, only further work will allow the determination of the proper scope and niche of DNA based computation and memory storage.

7. FUTURE WORK

Some centers of research in this area are at the University of Southern California at Los Angeles, with Dr. Adleman, Princeton, with Dr. Richard Lipton and his graduate students Dan Boneh and Eric Baum, and the NEC Research Institute in Princeton, NJ. With others elsewhere, they are developing new branches in this young field. Advancements are being made in cryptography. Researchers are working on decreasing error and damage to the DNA during the computations/reactions. The Princeton contingent has published papers on models for universal DNA computers, while others have described methods for doing addition and matrix multiplication with these computers.

Currently, molecular computing is a field with a great deal of potential, but few results of practical value. In the wake of Adleman's solution of the Hamiltonian path problem, there came a host of other articles on computation with DNA; however, most of them were purely theoretical. Currently, a functional DNA "computer" of the type most people are familiar with lies many years in the future. But work continues: in his article Speeding Up Computation via Molecular Biology Lipton shows how DNA can be used to construct a Turing machine, a universal computer capable of performing any calculation. While it currently exists only in theory, it's possible that in the years to come computers based on the work of Adleman, Lipton, and others will come to replace traditional silicon-based machines.

The field of DNA computing is truly exciting for the revolution it implies will occur within the next few years. It also demonstrates the current trend of merging and lack of distinction between the sciences, where a computer scientist can mess around with biology equipment and come up with something new and valuable.

8. REFERENCES

[1] R.W Keyes, "Silicon technology-in the chips for the future", Circuits and Devices Magazine, IEEE, 11,32-36 (1995)

[2] L. Adleman, "Molecular computations of solutions to combinatorial problems", Science 266, 1021-1024 (1994).

[3] S. Roweis and E. Winfree, "On the reduction of errors in DNA computation", Journal of Computational Biology, 6, 65-75 (1999).

[4] G. Păun, G. Rozenberg and A. Salomaa, "DNA Computing.New Computing Paradigms", Springer-Verlag, Berlin (1998)

[5] Gryko DT, Zhao, F, Yasseri AA, et al "Synthesis of thiol-derivatized ferrocene-porphyrins for studies of multibit information storage", J. Org. Chem, 65, 7356-7362 (2000)

[6] Roweis S, Winfree E, Burgoyne R, et al. "A sticker based model for DNA computation", J. Computational Bio, 5, 615-629 (1998)

[7] Macdonald Joanne, Stefanovic Darko and Stojanovic Milan N. "DNA computers for work and play", Scientific American, inc (2008)