An investigation of the role of DNA as molecular computers: A computational study on the Hamiltonian path problem

Alireza Heidari

Abstract — The ability of silicon to make faster computer chips ends soon. Thus, scientists have introduced DNA as a proper and clean alternative and source. Research on molecular computers is at initial steps. Molecular computers, in their current state, are composed of DNA and enzymes and there, still, remains a long way to achieving this dream. DNA computers will hopefully useful in various fields including treatment and complex computing.

Index Terms — DNA, Molecular computers, Enzymes, Biochemistry, Nucleic acids, Polymeric macromolecules, Biological molecules, Hamiltonian path problem, Traveling salesman problem, Nano drugs, Smart drugs

1 INTRODUCTION

Within the nucleus of every cell in our body, there is an intelligent microcomputer which is much more powerful than all the computers in the world. This tiny computer is in fact our DNA, or genetic material, which plan and sets all the information related to our live and function [1-4].

According to Moore’s Law, every 18 months, the number of electronic components on a computer chip doubles and the speed rises in many times [5]. However, electronic devices are getting smaller every day, but finally the physical speed and miniaturization for silicon microprocessors (semiconductors used in the manufacture of electronic circuits) will be approach to the end, so that the industrial production will be problematic [6, 7]. For years, chip makers have been looking for an alternative to replace silicon and this substitute is the DNA molecules present in the cells of living organisms; an abundant and inexpensive source which is an environmentally clean source, unlike to toxic conventional microprocessors [8].

2 DNA: MOLECULAR COMPUTERS

Today, a great revolution is taking place in the design and manufacture of computers and processors. Professionals have made interesting innovations based on the surrounding nature. The latest technology in this field is DNA computing that is designed and worked according to the genetics and DNA of living organisms.

Any genetic DNA molecule consists of four bases (C) cytosine, (A) adenine, (T) thymine, and (G) guanine that are entwisted together [9-16]. The language with which DNA acts is encoded by four letters of A, T, C, and G, and performs computations [17-21]. Using this four base system, the solution of each considered problem is encoded by using a DNA sequence put on the tape of a Turing machine [22-25]. Each stage can be created chemically in a test tube, by trillions of different DNA and then correct and acceptable stages are separated using genetic engineering tools [26]. Chemical bases of DNA molecules are biologically obeyed from certain rules, in a way that A base always bonds with T base and C always with G; these properties give processing capabilities to DNA molecule [27, 28]. For example, there is a molecule inside the cell with the size of 4 square nanometers, but can go down a DNA string and read any bases of A, T, C, G, made a complementary string of DNA from free afloat bases as the result, and then made the samples accessible [29].

3 RESULTS AND DISCUSSIONS

Everything was ready for Doctor Leonardo Edelman to bring his amazing idea; an old idea which originated by reading the book entitled as “Molecular Biology of the Gene” (by James Watson, who discovered DNA molecules in 1953). He concluded, based on the writings of Watson, that DNA molecules
have computing potential and now, this idea will be changed from a mental sparkle to a serious project in research centers [30]. Leonardo Edelman proved the ability of DNA molecules to solve complex math problems. The research results, published in the *Science* for the first time, astonished the world, so that calculations at the molecular level and with the help of the molecule of life, DNA, became the beginning of historical development in the field of computing and the emergence of a new generation of computers. Edelman using an interesting technique found a solution for one of the most famous computational problems, the Hamiltonian path problem, or the traveling salesman problem [31].

The problem is that a vendor was trying to choose the best path passing through a series of towns so that he must pass through each town only once and passes all the towns. Edelman solved this problem for the seven cities in question. However, the problem with the seven towns can be solved easily, even on paper, but in case the number of cities increases, it would be so difficult that will need several years to resolve it [32].

The importance of Edelman idea can be expressed as:

- Description of the possibility of using DNA molecules in solving classic math problems which solving them were impossible using conventional computational methods.
- Calculations carried out at the molecular level and eliminating obstacles in miniaturization of electronic components in the semiconductor industry.
- Proving unique aspects of DNA as a data structure.
- Proving that DNA solves the problems with the technique of parallel computation.

In this section, we explain the procedures of solving Hamiltonian problem using DNA molecules conducted by Leonard Edelman. Preparation of DNA sequences, which each of them represents a city and as we know, the genetic codes are displayed by letters. A number of rows composed of the four letters representing each city and the possible routes between the cities [33]. Then, these molecules are combined in a test tube and a number of these strings are attached to each other and a chain of these strands together represent one of the possible answers. Within seconds, all possible combinations of DNA strands which represent solutions are built in the test tube. Then, unacceptable molecules (e.g., the molecules which are not passed through some cities or passed twice a city) are eliminated through a series of chemical reactions. At the end, the strings that indicate the correct and acceptable paths remain in the test tube and are visible by microscope.

The problem solved by Edelman by DNA molecules just showed that although DNAs have the ability of complex computations, they cannot yet be used as substitutes for conventional electronic computers because they still need human help in their computations.

The aim of the researchers is a computer without human assistance. For example, in an experiment on Hamiltonian problem by Edelman, despite the fact that DNA molecules could very quickly create possible compounds which may be indicative of possible ways between cities, but he spent about a week in order to separate compounds which would not apply to answer the question through chemical reactions from the rest of the strings, and to achieve acceptable responses which is an indication of the need for human intervention in operations by the DNA.

Logic gates are one of the most important parts of a computer that can perform the operation that you command. These gates convert binary codes available in computer into a signal that computer uses to perform operations. Now, the logical gates receive input signals from silicon transistors, interpret them and then convert them into output signals which are used by computers to perform complex operations [34]. Production of logic gate consisting of DNA by the research team Rochester is the first step toward creating a computer with a structure similar to electronic computers. Instead of using electrical signals to perform logical operations, these gates act on DNA codes. They have identified a number of genetic materials as input and after connecting these components together they form a single output. For instance, a genetic gate called DNA connects two input DNAs using chemical laws and form a structure. Just as in the case of two Lego by a third logo are connected. Researchers believe that if these logic gates also are combined by microchips a great development will occur in manufacturing of DNA computers. DNA computers (logic gates, biochips) may take years to form a functional and usable DNA computer. But in case such a computer is made, scientists believe it would be much more compact, accurate, and efficient than today computers.

After the first test of Edelman on DNA molecules as a means of storing and processing information, i.e., solving the seven variable problem of Hamiltonian, his new trial is on the issue that there are over a million chances for the results, which solving it by humans are very complex without the help of computers. This problem needs to 20 values which creates a complex structure of communication [35].

A DNA computer acts in parallel with countless molecules that are communicate with each other in a few moments. While the time required to solve such complex problems on sequential computers exponentially increases (2, 4, 8, 16, …), this time only increases linearly (2, 4, 6, 8, …) in parallel computers. But now, Edelman explained that, unfortunately, DNA computers have some errors and their potential are
not in used to an acceptable manner. Edelman says “in the last century, we are well able to control electrons. But the use of DNAs in molecular computers will be a big change in technology”. He expresses that considering the computing and processing potential of DNAs, they will be used for non-biologic goals in addition to use in chemical and biological systems. He says “they are tiny molecular machines. They can save data and energy, attach and copy”. Edelman stated that “they have been made and growth during the past three million years, and we are just starting to use their potential for non-biological goals. Nature provided us an incredible toolbox and we began our search to find what will be possible to make by them” [36].

According to Moore’s Law, every 18 months, the number of electronic components on a computer chip doubles and the speed rises in many times. However, electronic devices are getting smaller every day, but finally the physical speed and miniaturization for silicon microprocessors (semiconductors used in the manufacture of electronic circuits) will be approach to the end, so that the industrial production will be problematic. For years, chip makers have been looking for an alternative to replace silicon and this substitute is the DNA molecules present in the cells of living organisms; an abundant and inexpensive source which is an environmentally clean source, unlike to toxic conventional microprocessors. Furthermore, according to the theory of Daniel Elie, DNA molecule conducts the electrons, as a molecular wire.

As noted above, Edelman, who is a mathematician and computer scientist, found the similarity between DNA function and computer systems by reading the Watson’s molecular biology book and studying the structure of DNA in Watson-Crick model [37].

DNA molecule is a double helix and to imagine the helix consider a backbone of alternating sugar and phosphate groups that wrapped around a long cylindrical projection; this is one of the strands of the double helix. There are two strands of the DNA helix holding each other by weak chemical bonds between bases that are outside of strands towards the molecules. In DNA, there are four types of nucleotides (structural unit of DNA) including Adenine (A), Thymine (T), Cystosine (C), and Guanine (G). According to this model, amounts of adenine and thymine are equal; because bases of Adenine in one of the two strands always bond to Thymine of opposite strand. Similarly, the amount of Guanine is equal to Cytosine since two bases in DNA are always bond together. Thus, if two strands of DNA molecule are separated by breaking the bonds between bases, each strand has the required data to synthesize of opposite strand. DNA self-replication is the ability that each genetic molecules must have as genetic material. DNA also is of this ability so that they self-replicated by separating of each two helix strands from each other and then making a new synthesized strand similar to the previous strand. More importantly, the Watson-Crick model indicated that genetic information is encoded in bases of DNA molecule just similar to the computers, i.e., saving data by binary strings consisting of two digits of 0 and 1. A binary digit called “bit”. Information on digital computers is displayed by a group of bits. Using coding techniques, bits are used to displaying discrete symbols like decimal numbers or alphabetic letters in addition to binary numbers. Using binary sets and implementing different coding methods, it can be possible to use sets of bites to construct complete set of computational codes. In topics of computer science, data are organized in different ways. Logical or mathematical model for a certain organization of data is called data structures. Data structures are in such a way that the data can be processed within them. Interestingly, an encoding DNA sequence by four bases of A, T, C, and G with a distance of about 0.35 nm between nucleotides, is a unique data structure.

On the other hand, the data concentration of DNA, or the capacity of DNA in saving data is incredible compared to the current computers. However, it can be possible to locate more than 10 trillion DNA molecules in a cubic centimeter. Such volume of DNA can store 10 Terabytes (1000 Gigabytes) of data and carry out 10 trillions of calculations in a moment. In addition, 1 gram of dried DNA, which is almost equal to a half of a sugar cube can store information in itself as a trillion CDs. Effective density of DNA is about 100000 times more than a modern hard drive.

Edelman with his talent realized that DNA in nature operates such as Turing machine. Turing machines called so in memory of English mathematician Alan Turing is an automaton; an abstract model of computer. Temporary memory of Turing machine is a tape. The tape is divided into cells, each able to maintain a mark. In connection with tape, there is a read-write head that can move left and right to read a sign in each move. Turing machine does not have a certain input file or output mechanism. Each input or output is done through the tape and having input and output files do not change the result. Based on the idea of Edelman and extensive researches of Professor Shapiro, DNA molecule such as Turing machine, processes the data and save them as a sequence or list of symbols. Hence, scientists have sought to create nano-computers by replacing silicon microprocessors with DNA molecules.

Nano-computer is a very tiny computer with very small circuits that can be seen only with a microscope. Nano-computers can be designed as follows:
(a) Electronic: Where nanotechnology is used to make microscopic circuits.
(b) Organic or biochemical: Similar to DNA computers.
Nano-computers are dealing with materials at the molecular dimensions and promising to build very small and faster computers. DNA computers are nano-scale computers which are composed of DNA molecules, enzymes to processing of molecules and a few drops of water. Input, output and software program of the machine, all are made of DNA molecules. In nature, there are enzymes that “cut” DNA molecules. In addition, there are enzymes that “paste”, “copy” and “repair” the DNA molecules. Furthermore, there is a unit like Central Processing Unit (CPU) of computers that operates some actions like “add”, “shifting bit” and logical operations (AND, OR, NOT, NOR), allows very complex computations. It is important to note that these actions only take a moment and implemented in parallel. Parallel processing means employing a variety of techniques in order to speed up the real-time data processing and analysis in computer systems. A parallel processing system is able to process the data simultaneously, rather than sequential processing, to more quickly process the data.

In other words, the objective of parallel processing is to speed up computer processing and increase the processing operation over a certain period of time.

In 2002, researchers at the Weizmann Institute of Science in Rehovot, led by Ehud Shapiro, unveiled a programmable molecular computing machine. This molecular computer were formed of the combination of enzymes and DNA molecules rather than silicon chips!

The molecular computer making team, succeeded to run the next step of their project after a year. In the new device, a single molecule of DNA that makes up the main part of the computer is able to produce enough energy to operate the computer through the data entry process. This project is considered a huge step in making DNA computers. Although the DNA computers are in early steps and their concept is in initial stages at now, they can alter the future of computers, especially in medical and biological applications of computers.

Biochemical nano computers are previously created in the nature; the existence of them in all organisms are already clear and evident. But they are uncontrollable by humans. For example, we cannot use a tree to calculate the number of pi.

Edelman, often known as the inventor of DNA computers. Since then, several research groups were proposed various schemes for DNA computers, but all of them are relied on a molecule called ATP as a source of energy [38-44]. But at the scheme proposed by Weizmann Institute, a redesign was performed so that the input data are act as a source of energy. In this device, the DNA should be considered as software and enzymes are used as hardware. Throw them together in a test tube. The procedure of chemical reactions between them allows to perform a simple two-operand.

Scientists have stated that these devices are worked by controlling the combination of software DNA molecules; this is a completely different process with what that occurred in current computers by the flow of electrons through a dry circuit. With the naked eye, a DNA computer looks like as a transparent solution of water in the test tube. In a DNA computer, there is no mechanical means. A drop of water is able to accommodate a trillion molecular biological tools. Instead of being displayed on the computer monitor screen, a technique that allows scientists to see the output length of DNA molecules is used to analyze the results. So far, DNA computers are only able to perform basic calculations and have not yet applied. Shapiro says “our computer is programmable, but it is not public and complete yet. There are still some computational tasks that DNA computers are not inherently capable of doing it”. The device can check whether a list of zero and one have an even number of one digits or not, but cannot count the number of one digits since its memory is limited and it is possible that the number of one digits are greater than its memory. These devices are also able to answer yes or no to a question, and as such, cannot correct a text with spelling mistakes. However, in terms of speed and size, DNA computers have surpassed the current computers. Just at the time that scientists argued about the incapability of silicon processors to more compaction, DNA molecules in nucleus of each molecule is able to save the data of more than one trillion music CDs, only in one cubic centimeter! While a personal computer can perform a calculation with high speed, a DNA computer, provides billions of possible responses simultaneously. This method causes the DNA computers have preference over current computers to solve problems of fuzzy logic, which have more possible solutions than binary logic in binary computers.

In the future, hybrid computers will be used. These computers are hybrids of traditional silicon chips for common computing tasks, and the DNA processor for specific computational tasks to take advantage of better things to do.

4 Future studies

Effective use of computers in the detection of lung and prostate cancers in the laboratory samples of Weizmann Institute and Ehud Shapiro's efforts as the team leader of computer experts and Biochemistry Institute, promising a bright future for the treatment of cancer and other unknown diseases. Soon, devices with molecular dimensions and with a combination of DNA molecules and required enzymes for detecting chemical signs of disease and then drug pumping to the desired point will be designed.

John Reif, who makes a personal molecular computer
in 2000, says in this regard: “In the future, it will be the case that doctor inject trillions of computers to the patient’s body.” They are designed so that focused on specific molecules to detect impairment, then they will be released other molecules that intervene the activity of cancer cells and destroy them. These computers are automatic and will not require chemical additives.

According to Shapiro, the injected molecules may interfere with the functioning of other cells or molecules, an issue that must be considered in advance because these interactions can be detrimental to the performance of the computers. The invention of nano drugs (smart drugs) is reported in the Nature. The duty of these drugs is to fix the illness, in desired position, by relieving the biological molecules to make changes in considered cells.

In any case, the emergence of nanoscale is not a dream and is not like to the movie “dream trip” (1996) in which a surgical team miniaturized by a special device and entered into the body of a patient. Scientists are believed that the complexity of structure of biological molecules allows that the DNA computers perform their tasks without need to electronic parts.

As already mentioned, in a DNA computer, DNA molecules and enzymes are used as input, output, software, and hardware. In 2003, a power source also added to these computers. The source provides its energy from breaking the DNA molecules. In mid-August 2003, another group of scientists stated in an article that they made a biological computer that is designed for gaming Toe Tac Tic. They believe that the computer never loses the game to human. Another use of DNA computers is the expertise in the analysis of gene expression and computing DNA compounds, chemical reactions enforcement, search, and extracting desired results.

DNA computers open wonderful gates to the unknown combination of genetics and engineering and promise days with no trace of incurable diseases. Tools and nanomedicine are becoming the alternative of medical examinations, treatment will do easily and every man can be a medical doctor. Astronomy and space calculations are done in less than a minute. The advent of DNA computers means a historical development in the computer industry and the emergence of a new generation of computers; computers that are very different in size and shape from today’s conventional computers. Although such computers would not be a relevant tool to gaming or run a MP3 file, surely they are capable to solve the logical, security and hiding problems. The next step in the use of DNA computers is studying the genetic programming, algorithms, automata, and designing the system languages.

5 CONCLUSIONS

Growth of human knowledge is always led to looking for new resources. In the field of computer manufacturing industry, human found a lot. Once, it appeared that silicon was a good source for making computer chips, but nowadays, the end of silicon power has observed by scientists. So, scientists are trying to keep the process of scientific development in this field. It appears that DNA can be a good natural source in this field for a long time. Low volume is one of the most important features of this source. DNA computers are not distant dreams but are achievable reality. Computers that are smaller than a water droplet and a trillion of them are fit in a test tube.

REFERENCES

[12] Inken Wierstra, Chapter Three - The Transcription Factor FOXM1 (Forkhead box M1): Proliferation-Specific Expression, Transcription Factor Function, Target Genes, Mouse Models, and Normal Biological Roles, In: