



A
Seminar Report
On
“DNA Computing”

Submitted in partial fulfillment of
the Bachelor of Engineering Degree
of the University of Rajasthan, Jaipur



LAXMI DEVI INSTITUTE OF
ENGINEERING & TECHNOLOGY
CHIKANI ALWAR (RAJASTHAN)

Session: 2005-2006

Submitted to
Mr Ajay Khunteta

HOD CE Deptt.
LIET, Alwar

Guided by:
Mr Vikas Thada

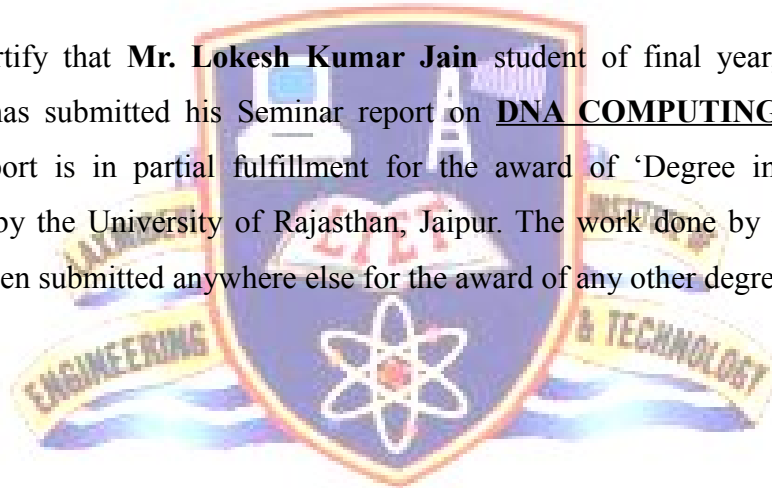
Lect. CE Deptt.
LIET, Alwar

Submitted By
Lokesh Kumar Jain
VIII Semester CE



CERTIFICATE

This is to certify that **Mr. Lokesh Kumar Jain** student of final year, COMPUTER Engineering has submitted his Seminar report on **DNA COMPUTING**. The seminar work and report is in partial fulfillment for the award of 'Degree in COMPUTER Engineering' by the University of Rajasthan, Jaipur. The work done by him is genuine and has not been submitted anywhere else for the award of any other degree or diploma.



विद्या ददाति विनयम्

Submitted to
Mr Ajay Khunteta
HOD CE Deptt.
LIET, Alwar

Guided By
Mr Vikas Thada
Lect. C.E. Deptt.
LIET, Alwar



ACKNOWLEDGEMENT

I am greatly thankful to my seminar guide **Mr Vikas Thada** Computer Engineering Department, who inspired me to present my seminar on **"DNA COMPUTING"**.

He helped and encouraged me in every possible way. The knowledge acquired during the preparation of the seminar report would definitely help me in my future ventures.

I would like to express my sincere gratitude to Mr Vikas Thada, Lecturer, Department of *computer Engineering*, for finding out time and helping me in this seminar.

विद्या ददाति विनयम्

I would also thank all the teachers of our Department for their help in various aspects during the seminar.

Date:

Lokesh Kumar Jain
VIII Semester
Computer Engineering



INDEX

- DNA
- DNA STRUCTURE
- INTERESTING FACTS
- WHAT IS NEEDED?
- WHERE IT ALL STARTED?
- HOW IT WORKS?
- DNA CHIP
- ADVANTAGES
- CHALLENGES TO IMPLEMENTATION
- GOALS FOR THIS WORK
- APPLICATIONS
- LIMITATIONS
- LATEST DEVELOPMENTS
- COMPARISON OF DNA COMPUTERS WITH CONVENTIONAL COMPUTER
- FEATURES OF DNA COMPUTER
- DNA BASICS
- OPERATIONS ON DNA SEQUENCES



DNA Computing **(Deoxyribonucleic Acid Computing):**

DNA computing is a nascent technology that seeks to capitalize on the enormous informational capacity of DNA, biological molecules that can store huge amounts of information and are able to perform operations similar to a computer's through the deployment of enzymes, biological catalysts that act like software to execute desired operations.

Scientists around the globe are now trying to marry computer technology and biology by using nature's own design to process information. Research in this area began with an experiment by Leonard Adleman, a computer scientist at USC who surprised the scientific community in 1994 by using the tools of molecular biology to solve a hard computational problem.

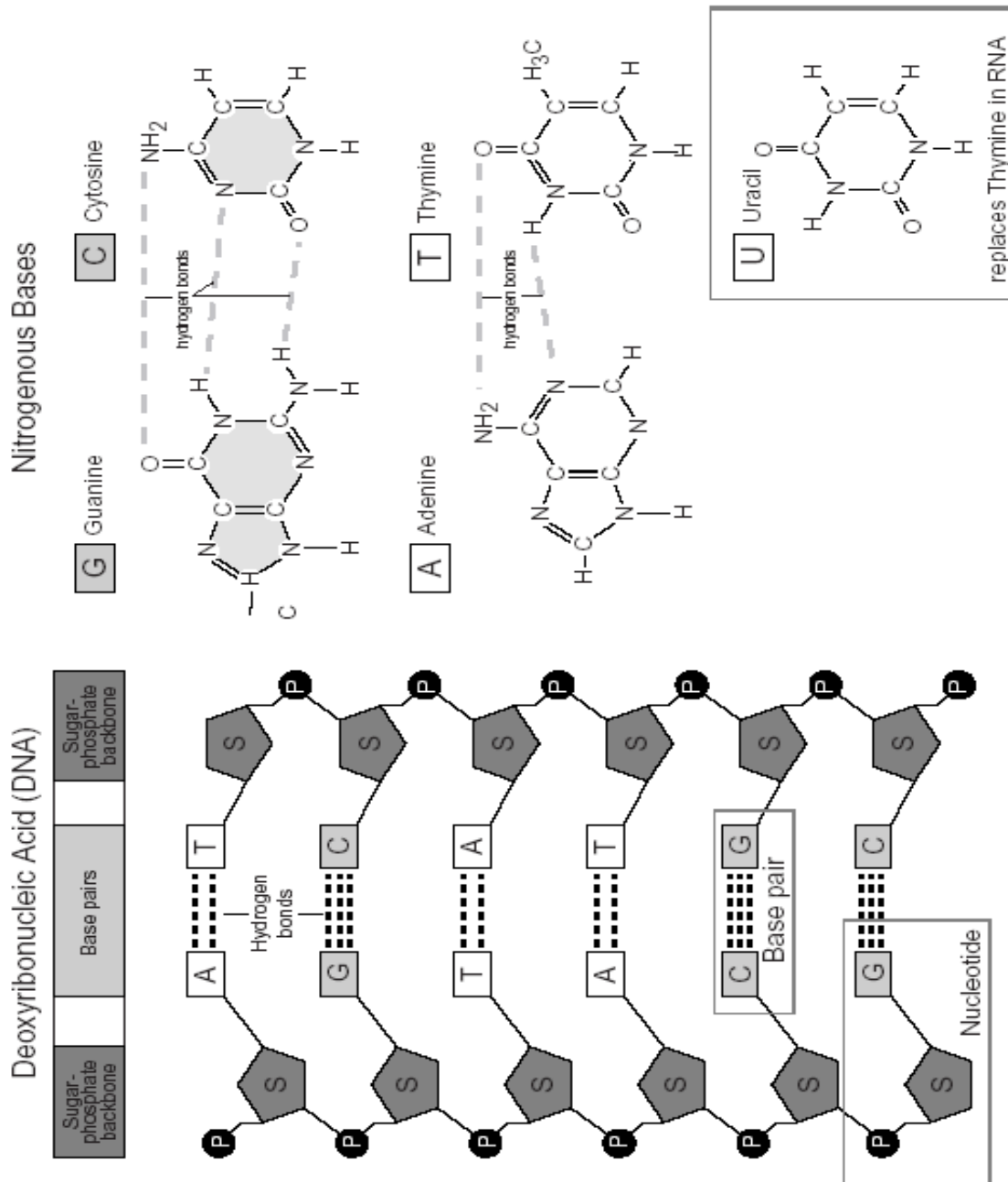
A new version of a biomolecular computer developed at the Israel Institute of Technology composed entirely of DNA molecules and enzymes. It can perform as many as a billion different programs simultaneously. Previous biomolecular computers, such as the one built by Institute of Science three years ago, were limited to just 765 simultaneous programs.

This new computer is also autonomous; it processes calculations from beginning to end without any human assistance. Other biomolecular computers require humans to analyze and decipher results and perform intermediate tasks at different points in the process before the computer can complete the operation.

Current computers consist of metal, plastic, wires and transistors. The manner in which they process information is called linear because they conduct one computation at a time. In the latest generation of computers, biological molecules replace all the components. One advantage of these biomolecular computers over linear computers is their ability to simultaneously carry out an enormous number of complex operations.



DNA Structure:





Interesting Facts:

- DNA molecule is 1.7 meters long
- Stretch out the entire DNA in your cells and you could reach the moon 6000 times!
- DNA is the basic medium of information storage for all living cells. It has contained and transmitted the data of life for billions of years
- Roughly 10 trillion DNA molecules could fit into a space the size of a marble. Since all these molecules can process data simultaneously, you could theoretically have 10 trillion calculations going on in a small space at once.



DNA Lab Chip



DNA Molecule



WHAT IS THE NEED?

Computers have become significantly smaller and more powerful over the past 40 years, but they still have a silicon substrate, and silicon has inherent limitations. The abilities and power of computers to this day have increased, almost exponentially, since the dawn of their creation. This exponential growth of silicon chip speed and inverse of size has come to be known as Moore's Law. Computer chip manufacturers are furiously racing to make the next microprocessor that will topple speed records. As advancements in micro silicon chip production continue, however, more and more obstacles are faced due to the increase in complexities of the problems for which they are required. Chip makers need a new material to produce faster computing speeds.

It would be hard to believe where scientists have found the new material they need to build the next generation of microprocessors. Millions of natural supercomputers exist inside living organisms, including our body. DNA (deoxyribonucleic acid) molecules, the material our genes are made of, have the potential to perform calculations many times faster than the world's most powerful human-built computers. DNA molecules have already been harnessed to perform complex mathematical problems. The fastest supercomputers now available can perform about 10^9 (1 billion) operations per second. By using DNA molecules, it would be possible to achieve effective speeds of as much as 10^{17} operations per second.

WHERE IT ALL STARTED?

The scientists at the forefront of the DNA computer revolution are a brilliant breed indeed. It was all started by a professor of Computer Science at USC by the name of Leonard M. Adleman, who utilized recombinant DNA to solve a simple Hamiltonian path problem, more popularly recognized as a variant of the so-called "traveling salesman problem." In Adleman's version of the traveling salesman problem, or "TSP" for short, a hypothetical salesman tries to find a route through a set of cities so that he visits each city *only once*. As the number of cities increases, the problem becomes more difficult until its solution is beyond analytical analysis altogether, at which point The Hamiltonian path problem, on a large scale, is effectively unsolvable by conventional computer systems. Computers now solve such problems by trial and error. But if hundreds of cities were involved, a conventional computer would require years to find

DNA Computing



the answer. A DNA computer, on the other hand, tests all possible answers simultaneously, offering the prospect of much speedier solutions.

HOW IT WORKS?

DNA computation is based on the fact that technology allows us to 'sequence' (design) single DNA strands which can be used as representations of bits of binary data. Technology also allows us to massively 'amplify' (reproduce) individual strands until there are sufficient numbers to solve complex computational problems.

- DNA input molecule
- The famous double-helix structure discovered by Watson and Crick consists of two strands of DNA wound around each other. Each strand has a long polymer backbone built from repeating sugar molecules and phosphate groups. Each sugar group is attached to one of four "bases". These four bases - guanine (G), cytosine (C), adenine (A) and thymine (T) - form the genetic alphabet of the DNA, and their order or "sequence" along the molecule constitutes the genetic code.



Generic Code

In the cell, DNA is modified biochemically by a variety of enzymes, which are tiny protein machines that read and process DNA according to nature's design. Just like a CPU has a basic suite of operations like addition, bit-shifting, logical operators (AND, OR, NOT NOR), etc. that allow it to perform even the most complex calculations; DNA has cutting, copying, pasting, repairing, and many others. Many copies of the enzyme can work on many DNA molecules simultaneously. This is the power of DNA computing, that it can work in a massively parallel fashion. Pairs of molecules on a strand of DNA represent data and two naturally occurring enzymes act as the hardware to read copy and manipulate the code.

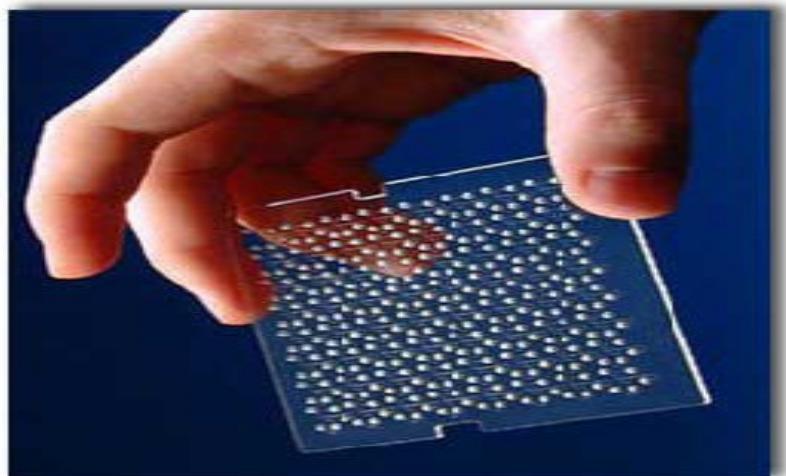
DNA Computing



DNA Chip:



Mother Board





DNA molecule Arrangement in Chip

ADVANTAGES:

DNA computers derive their potential advantage over conventional computers from their ability to:

- Perform millions of operations simultaneously. The massively parallel processing capabilities of DNA computers may give them the potential to find tractable solutions to otherwise intractable problems, as well as potentially speeding up large, but otherwise solvable, polynomial time problems requiring relatively few operations.
- Another advantage of the DNA approach is that it works in "parallel," processing all possible answers simultaneously. Therefore it enables to conduct large parallel searches and generate a complete set of potential solutions.
- DNA can hold more information in a cubic centimeter than a trillion CDs, thereby enabling it to efficiently handle massive amounts of working memory.
- The DNA computer also has very low energy consumption, so if it is put inside the cell it would not require much energy to work and its energy-efficiency is more than a million times that of a PC.

Challenges to Implementation:

- Practical protocols for input and output of data into the memory.
- A Representation of data in DNA sequences.
- An Understand the information capacity of the hybridization interactions in large collections of many different DNA sequences.
- Appropriate physical models to guide design and experimentation

Goals for This Work:

- Simplicity in design and practice.
- A Learn DNA sequences to which the memory is exposed, and capture contextual sequence information.
- A Use hybridization affinity for associative recall, and generalization to new input through sequence similarity.

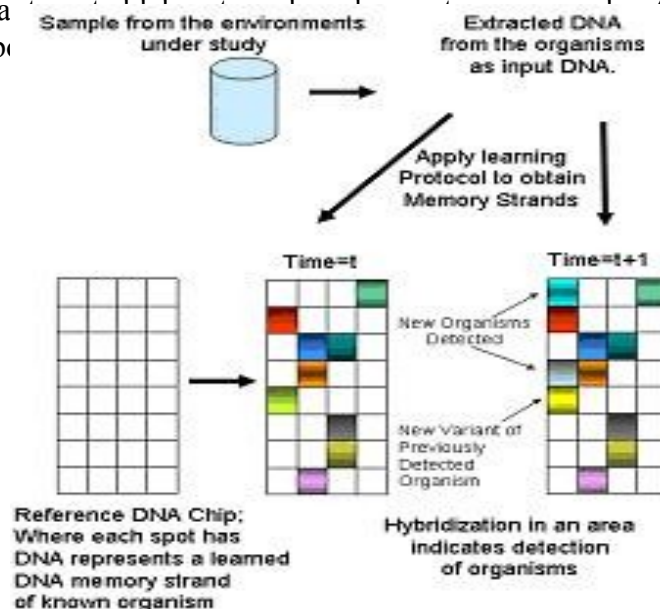


- A Use Non-Cross hybridizing Tag system to decouple IO from specific sequences in the memory.

APPLICATIONS:

The potential applications of re-coding natural DNA into a computable form are many and include:

- DNA sequencing
- DNA fingerprinting
- DNA mutation detection
- Development and miniaturization of biosensors, which could potentially allow communication between molecular sensory computers and conventional electronic computers.
- The fabrication of nanoscale objects that can be placed in intracellular locations for monitoring and modifying cell function
- The replacement of silicon devices with nanoscale molecular-based computational systems, and The application of biopolymers in the formation of novel nanostructured materials with unique optical and selective transport properties
- DNA based models of computation might be useful for simulating or modeling other emerging computational paradigms, such as quantum computing, which may not be feasible until much later.
- Evolutionary programming for applications in design or expert systems.
- In theory, this technology could one day lead to the development of hybrid computer systems, in which a silicon-based PC generates the code for a 'lab in a bi





LIMITATIONS:

However, there are certain shortcomings to the development of the DNA computers:

- A factor that places limits on this method is the error rate for each operation. Since these operations are not deterministic but stochastically driven, each step contains statistical errors, limiting the number of iterations one can do successively before the probability of producing an error becomes greater than producing the correct result.
- Algorithms proposed so far use relatively slow molecular-biological operations. Each primitive operation takes hours when you run them with a small test tube of DNA. Some concrete algorithms are just for solving some concrete problems. Every Generating solution sets, even for some relatively simple problems, may require impractically large amounts of memory. Also, with each DNA molecule acting as a separate processor, there are problems with transmitting information from one molecule to another that have yet to be solved.



LATEST DEVELOPMENTS:

Israeli scientists have devised a computer that is so tiny that a trillion of them could fit in a test tube and perform can perform 330 trillion operations per second, more than 100,000 times the speed of the fastest PC with 99.8 percent accuracy. It is the first programmable autonomous computing machine in which the input, output, software and hardware are all made of biomolecules. Recently, the team has gone one step further. In the new device, the single DNA molecule that provides the computer with the input data also provides all the necessary fuel.

Classical DNA computing techniques have already been theoretically applied to a real life problem: breaking the Data Encryption Standard, DES. Although this problem has already been solved using conventional techniques in a much shorter time than proposed by the DNA methods, the DNA models are much more flexible, potent, and cost effective.

Israeli scientists have devised a computer composed of DNA and enzymes. The enzyme FokI breaks bonds in the DNA double helix, causing the release of enough energy for the system to be self-sufficient. The design is considered a giant step in DNA computing which could transform the future of computers, especially in pharmaceutical and biomedical applications.





FIRST DNA COMPUTER

Olympus Optical Co. – First practical DNA Computer specification

- Tokyo (July 3rd, 2002)
- Olympus Optical Co. Ltd.
- First commercially practical DNA computer
- Specializes in gene analysis
- Akira Toyama, an assistant professor at Tokyo University.
- Standard gene analysis approach very time consuming (3 days)
- Now done in 6hrs
- Joint project called NovousGene Inc. spec in genome informatics.
- Available for commercial use by researchers by 2003 sometime.
- Two sections
 - Molecular Calculation component
 - DNA combination of molecules
 - Implements chemical reactions
 - Searches
 - Pulls out right DNA results
 - Electronic Calculation component
 - Executes processing programs
 - Analysis these results

—

Comparison of DNA computers with conventional Computer:

Computing with DNA is a completely new method

DNA Computing



among the quantum computing. Alternative to electronic/semiconductor technology, computing with DNA use biochemical process based on DNA. Computing with DNA is also known as *molecular computing*, a new approach to massive parallel computation based on groundbreaking work by *Leonard Adleman*.

DNA plays the role of information storage in nature. DNA is the genetic material containing the whole information of an organism to be copied into the next generation of the species. DNA computing is a computational paradigm that uses synthetic (or natural) DNA molecules as information storage media. The techniques of molecular biology, such as polymerase chain reaction (PCR), gel electrophoresis, and enzymatic reactions, are used as computational operators for copying, sorting, and splitting/concatenating the information in the DNA molecules, respectively.

Computing with DNA molecules has many advantages over conventional computing methods that utilize solid-state semiconductors. The properties of DNA computing compared with conventional computers are summarized in Table 1. Though DNA computing performs individual operations slowly, it can execute billions of operations simultaneously. This is contrasted with the electronic digital computers where individual operations are very fast; however, the operations are executed basically sequentially. The massive parallelism of DNA computing comes from the huge number of molecules, which chemically interact, in a small volume. DNA also provides a high storage capacity since they encode information on the molecular scale.

<u>Basics</u>	<u>DNA Computers</u>	<u>Conventional Computers</u>
Storage Media	Nucleic acids	Semiconductors
Memory Capacity	Ultra-High	High
Operators	Biochemical Operations	Logical Operations (and, or, not)
Operations	Simultaneous (Parallel)	Bitwise (Sequential)
Speed of each Operation	Slow	Fast
Process	Stochastic	Deterministic

Features of DNA computer:

- Storage capacity: The information density could go up to 1 bit/nm^3 .
- High parallelism: every molecule could act as a small processor on nano-scale and the number of such processors per volume would be potentially enormous. In an in vitro assay we could handle easily with



about 10^{18} processors working in parallel.

- Speed: Although the elementary operations (electrophoresis separation, ligation, and PCR-amplifications) would be slow compared to electronic computers, their parallelism would strongly prevail, so that in certain models the number of operations per second could be of order 10^{18} operations per second, which is at least 100,000 times faster than the fastest supercomputers existing today.
- Energy efficiency: It performs 10^{19} operations per Joule. This is about a billion times more energy efficient than today's electronic devices.

DNA BASICS:

DNA encodes the genetic information of cellular organisms. It consists of polymer chains, commonly referred to as DNA strands. Each strand may be viewed as a chain of nucleotides, or bases. An n -letter sequence of consecutive bases is known as an n -mer or an oligonucleotide of length n . The four DNA nucleotides are adenine, guanine, cytosine and thymine, commonly abbreviated to A, G, C and T respectively. Each strand has, according to chemical convention, a 5' and a 3' end, thus any single strand has a natural orientation. The classical double helix of DNA is formed when two separate strands bond together. Bonding occurs due to complementary base pairing. Bonds with T and G bonds with C. The pairs (A, T) and (G, C) are therefore known as Watson-Crick complementary base pairs. Thus a hypothetical DNA molecule sequence is

5'-ACGCGTACGTACAAGTGTCCGAAT-3'
3'-TGC GCATGCATGTTACAGGCTTA-5'

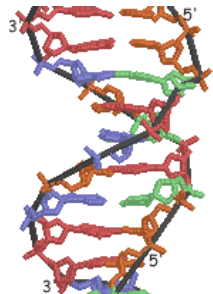


Fig. The classical double helix of DNA showing Watson-Crick complementary base pairs (A, T, G, C).

- Deoxyadenylate (A) is in blue
- Deoxythymidylate (T) is in green

DNA Computing

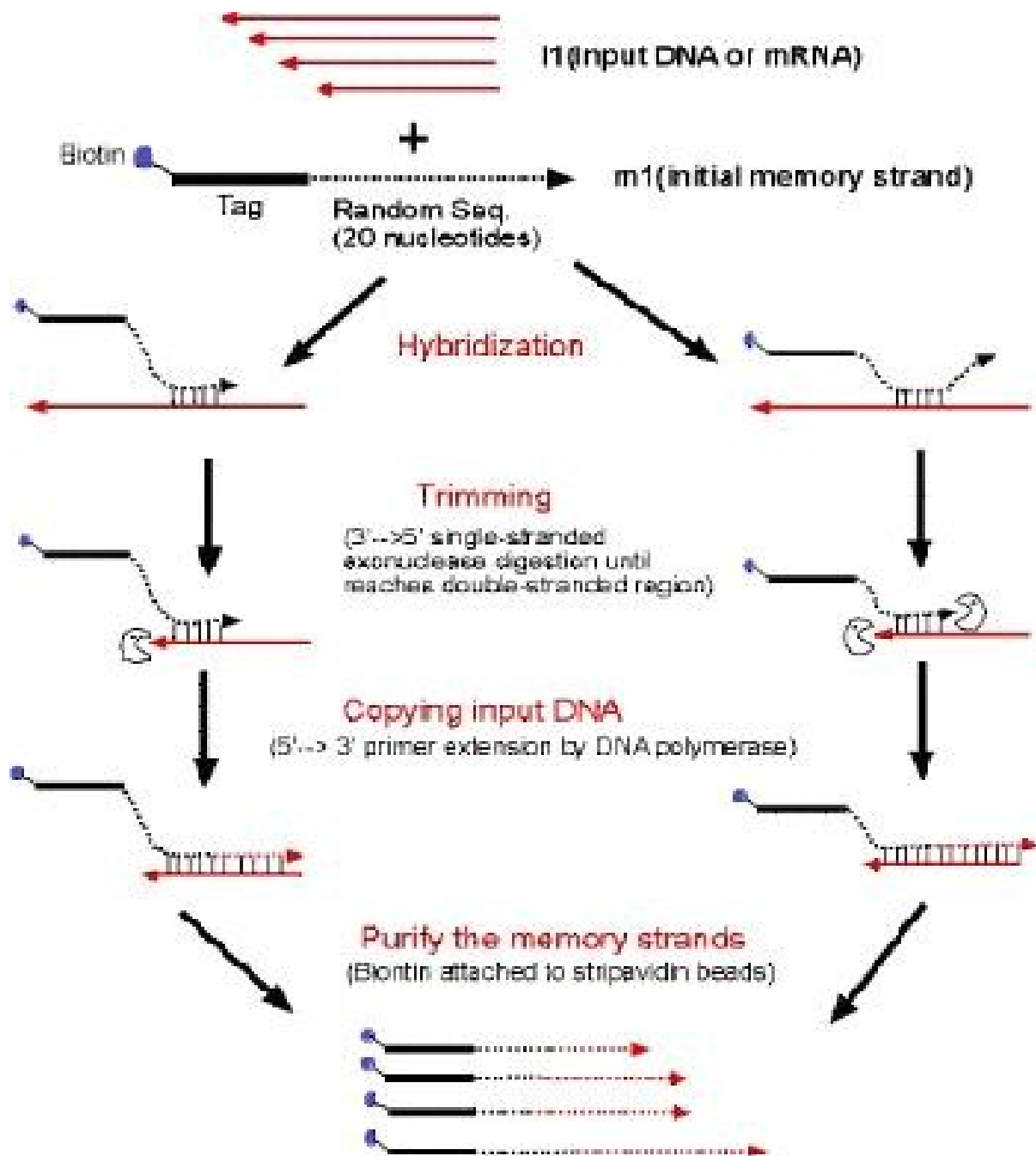


- Deoxyguanylate (G) is in red
- Deoxycytidylate (C) is in orange

Operations on DNA sequences:

The following operations can be done on DNA sequences in a test tube to program the DNA computer:

- Synthesis: synthesis of a desired strand
- Separation: separation of strands by length
- Merging: pour two test tubes into one to perform union
- Extraction: extract those strands containing a given pattern
- Melting/Annealing: break/bond two single strand DNA molecules with complementary sequences. Amplification: use PCR to make copies of DNA strands
- Cutting: cut DNA with restriction enzymes
- Ligation: Ligate DNA strands with complementary sticky ends using Ligase.
- Detection: Confirm presence/absence of DNA in a given test tube.
- Binding: Cooling of single strand solution below $85^{\circ} - 95^{\circ} \text{C}$ makes strands fuse again.
- Multiplying: Produces 2^n copies of double stranded DNA sequence α ..



memory strands are ssDNA each with a unique tag plus various length of 3' regions that are complementary to the input DNA or RNA



Bibliography

1. <http://www.usc.edu/dept/molecular-science/papers/adleman-science.pdf>
2. WWW.ieee.org
3. www.acm.org
4. Adleman Original Papers
5. www.google.com