DNA Computers

Dr. Shyam Kamal

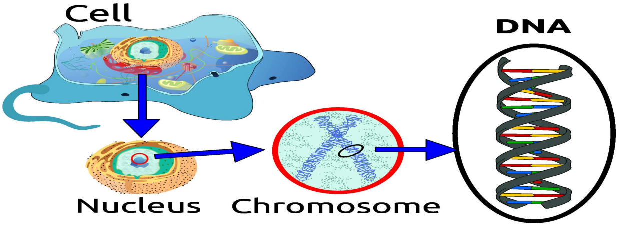
Department of Systems Design and Informatics
Kyushu Institute of Technology Japan

kamal@ces.kyutech.ac.jp

13 June 2016

Overview

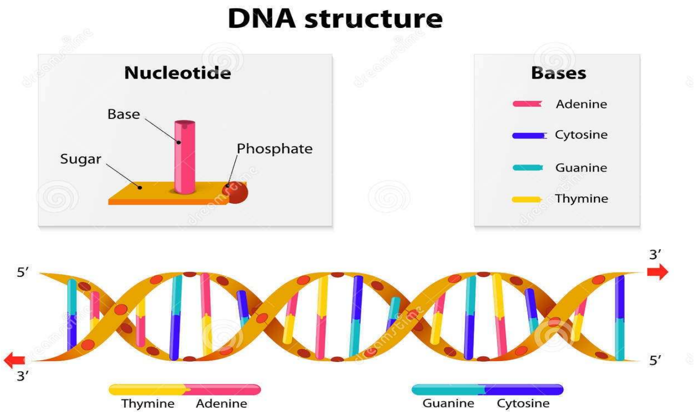
Genetic Information



Chromosomes

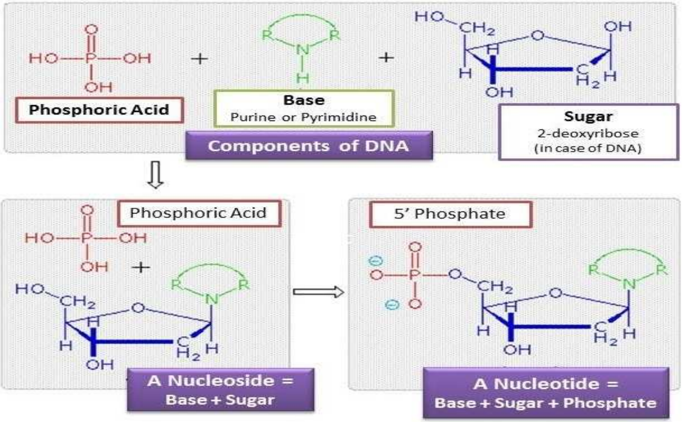
- composed of protein and deoxyribonucleic acid (DNA)
- both parents contribute material that determines the characteristics of their offspring
- DNA carries the genetic information

DNA strands



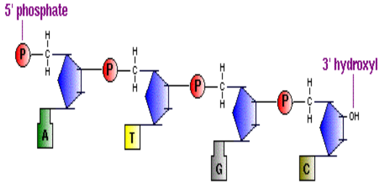
- DNA consists of polymer chains, referred to as DNA strands.

DNA strands in more detail

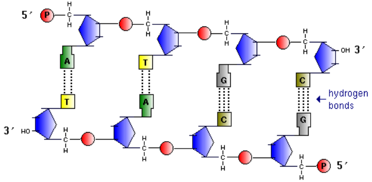


- Any single strand has a natural orientation : a free 5‘ phosphate group, and the other end has a free 3‘ deoxyribose hydroxl group.

Single and Double Strand



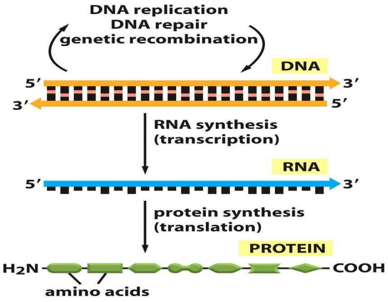
(a)



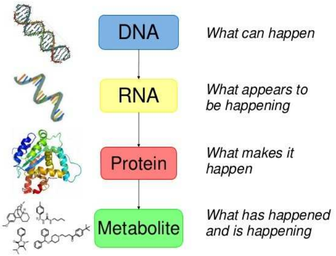
(b)

- Bonding occurs by the pairwise attraction of bases; A bonds with T and G bonds with C.
- The pairs (A,T) and (G,C) are therefore known as complementary base pairs.
- The two pairs of bases form hydrogen bonds between each other

Transcription and Translation



(c)



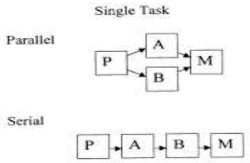
(d)

DNA supports two key functions for life

- (1) Coding for the production of proteins which are the working molecules in organisms, (2) Self-replication.
- Remark: One of the most important functions proteins carry out is to act as enzymes.

Reasons for using DNA based Computation

(e)



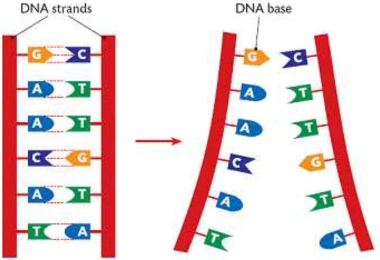
(f)

- The information density of DNA is much greater than that of silicon: 1 bit can be stored in approximately one cubic nanometer. Others storage media, such as videotapes, can store 1 bit in 1,000,000,000,000 cubic nanometer.
- Operations on DNA are massively parallel: a test tube of DNA can contain trillions of strands. Each operation on a test tube of DNA is carried out on all strands in the tube in parallel.

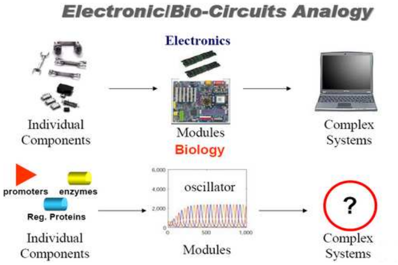
Principles of DNA Computing

- **Duplication and Transmission of Information:**
  - DNA can either duplicate the information in DNA molecules or transmit this information to other DNA molecules.
- **Representation of Information:**
  - Present computer using electrical impulses to represent bits of information,
  - the DNA computer uses the chemical properties of these molecules by examining the patterns of combination or growth of the molecules or strings.
- **Execution of the desired Calculation:**
  - DNA can do this through the manufacture of enzymes,
  - which are biological catalysts that could be called the 'software', used to execute the desired calculation.

Analogy of Electronic Computer



(g)



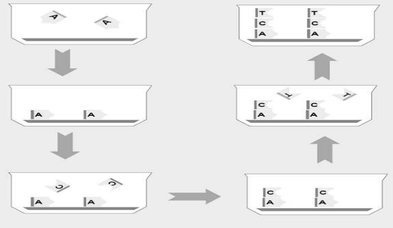
(h)

- Electronic computer encode information using $\{0\ 1\}$, similarly DNA using $\{A\ G\ C\ T\}$
- 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.

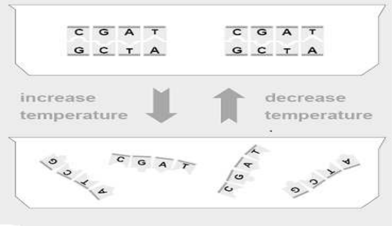
Why DNA is a convenient choice?

- **DNA is a convenient choice because**
  - it is both self-complementary, allowing single-stranded DNA to select its own complement
  - can easily be copied
- **Molecular biologists have already built a toolbox for manipulating DNA, including**
  - restriction enzyme digestion
  - ligation
  - sequencing
  - amplification
  - fluorescent labelling

Bio-operations: Synthesizing, Annealing and Melting



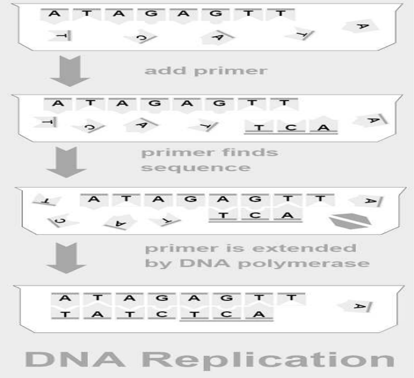
(i)



(j)

- **Synthesizing:** a desired polynomial-length strand used in all models.
- **Annealing:** bond together two single-stranded complementary DNA sequences by cooling the solution. Annealing in vitro is known as hybridization.
- **Melting:** break apart a double-stranded DNA into its single-stranded complementary components by heating the solution. Melting in vitro is also known under the name of denaturation.

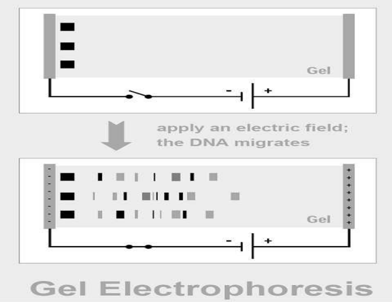
DNA Amplifying:



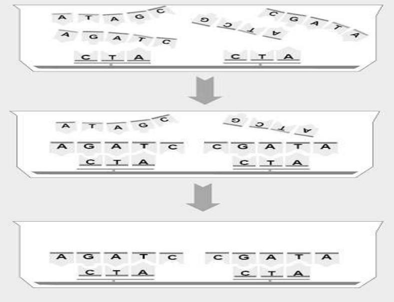
Amplifying (copying):

- make copies of DNA strands by using the Polymerase Chain Reaction.
- The replication reaction requires a guiding DNA single-strand called template, and a shorter oligonucleotide called a primer, that is annealed to it.

Bio-operations: Separating and Affinity purification



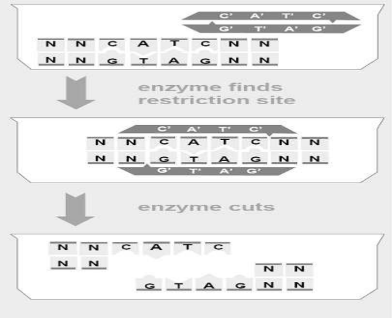
(k)



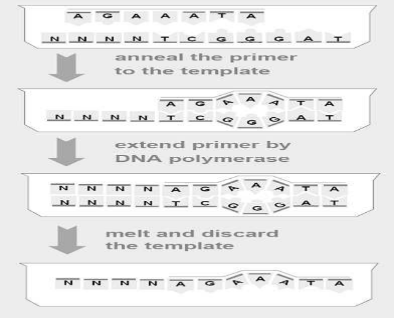
(l)

- **Separating :** the strands by length using a technique called gel electrophoresis that makes possible the separation of strands by length.
- **Extracting :** those strands that contain a given pattern as a substring by using affinity purification.

Bio-operations: Cutting and Substituting



(m)



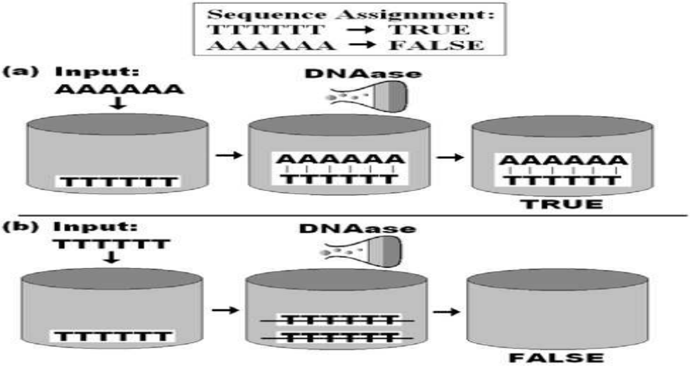
(n)

- **Cutting :** DNA double-strands at specific sites by using commercially available restriction enzymes.
- **Substituting:** substitute, insert or delete DNA sequences by using PCR site-specific oligonucleotide mutagenesis.

Some other Bio-operations

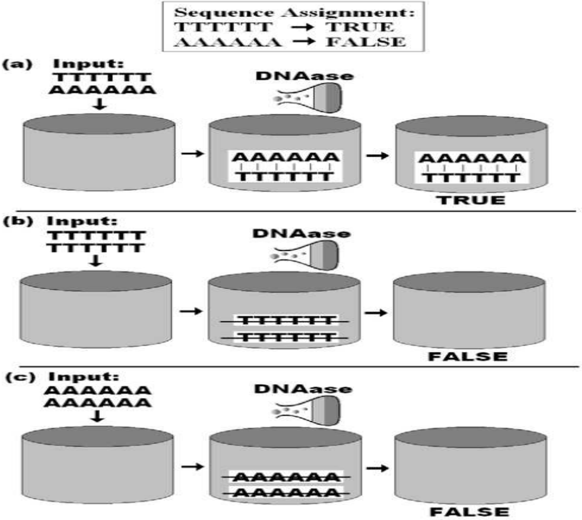
- **Ligating:**
  - paste DNA strands with compatible sticky ends by using DNA ligases.
  - Indeed, another enzyme called DNA ligase, will bond together, or “ligate”, the end of a DNA strand to another strand.
- **Marking single strands by hybridization:**
  - complementary sequences are attached to the strands, making them double-stranded.
  - The reverse operation is unmarking of the double-strands by denaturing, that is, by detaching the complementary strands.
- **Destroying the marked strands by using exonucleases:**
  - By using enzymes called exonucleases, either double-stranded or single-stranded DNA molecules may be selectively destroyed.
- **Detecting and Reading:**
  - given the contents of a tube, say “yes” if it contains at least one DNA strand, and “no” otherwise.
  - PCR may be used to amplify the result and then a process called sequencing is used to actually read the solution.

DNA-Based NOT Gate

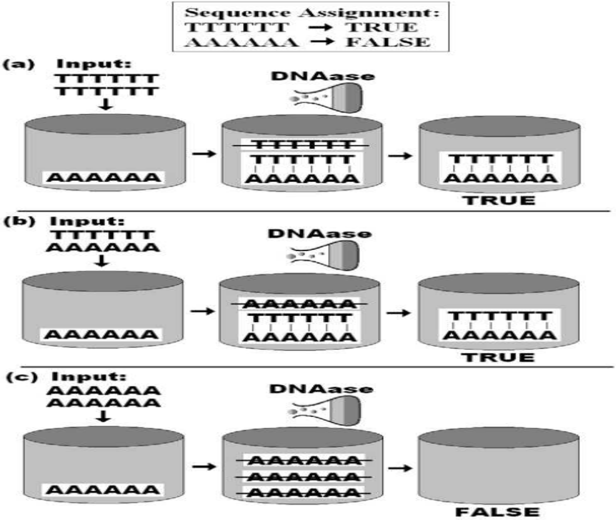


- (a) A false input results in a double stranded sequence representing a truth evaluation
- (b) a true input does not the input sequence is true, then TTTTTT will not bind with the base TTTTTT sequence. DNAase will destroy both sequences; no double stranded sequences will be observed.

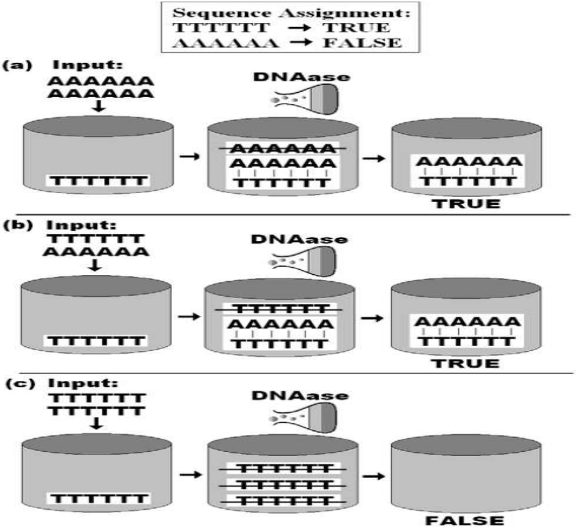
DNA-Based XOR Gate



DNA-Based OR Gate



DNA-Based NAND Gate



Concluding Remarks

Advantages

- **Perform millions of operations simultaneously.**
- **Generate a complete set of potential solutions and conduct large parallel searches.**
- Efficiently handle massive amounts of working memory.
- They are inexpensive to build, being made of common biological materials.
- **The clear advantage is that we have a distinct memory block that encodes bits.**
- Using one template strand as a memory block also allows us to use its compliment as another memory block, thus effectively doubling our capacity to store information.

Concluding Remarks Continued...

Disadvantages:

- **Generating solution sets, even for some relatively simple problems, may require impractically large amounts of memory (lots and lots of DNA strands are required).**
- **Many empirical uncertainties, including those involving: actual error rates, the generation of optimal encoding techniques, and the ability to perform necessary bio-operations conveniently in vitro.**

DNA computers could not (at this point) replace traditional computers.

The End