Lecture 1

# Introduction:
# The Molecular Basis of Life

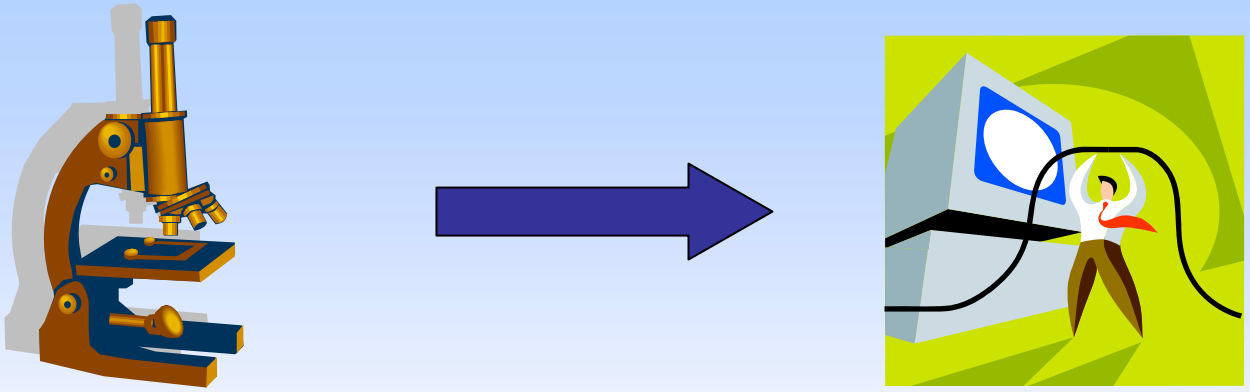
# Life

- A living organism is *an open, self-regulating and self-replicating system built from bio-molecules*

Visit: **DNA from the Beginning**

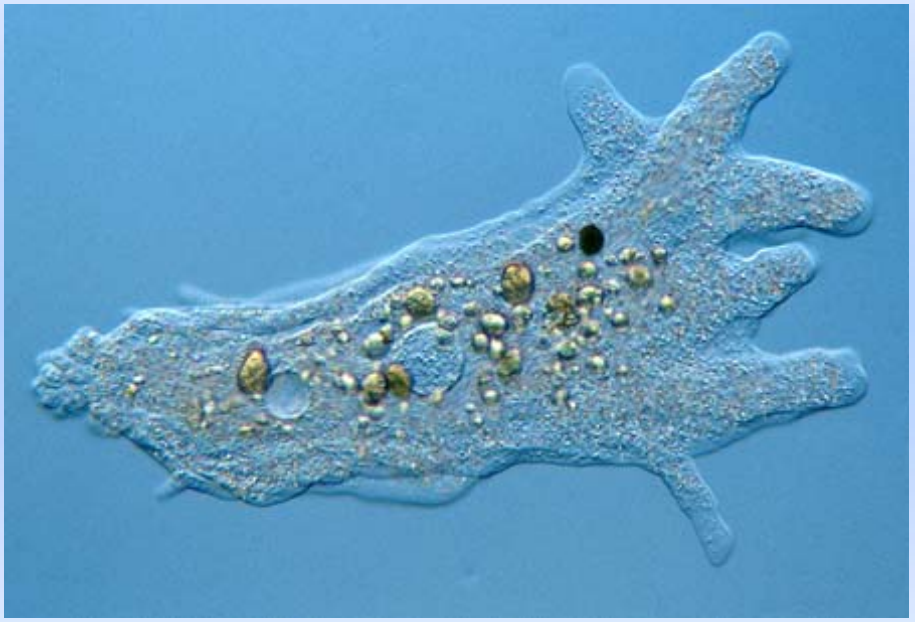
# An Historical Perspective

- … – 1900 **Pre-Mendelian period**
- 1900 – 1940 Pre-DNA period
- 1940 – 1990 DNA period
- 1990 – 2003 Genomic period
- 2003 – … Post-genomic era



# The Pre-Mendelian Period

- Vital force
- Cells



1670s -Unicellular - amoeba

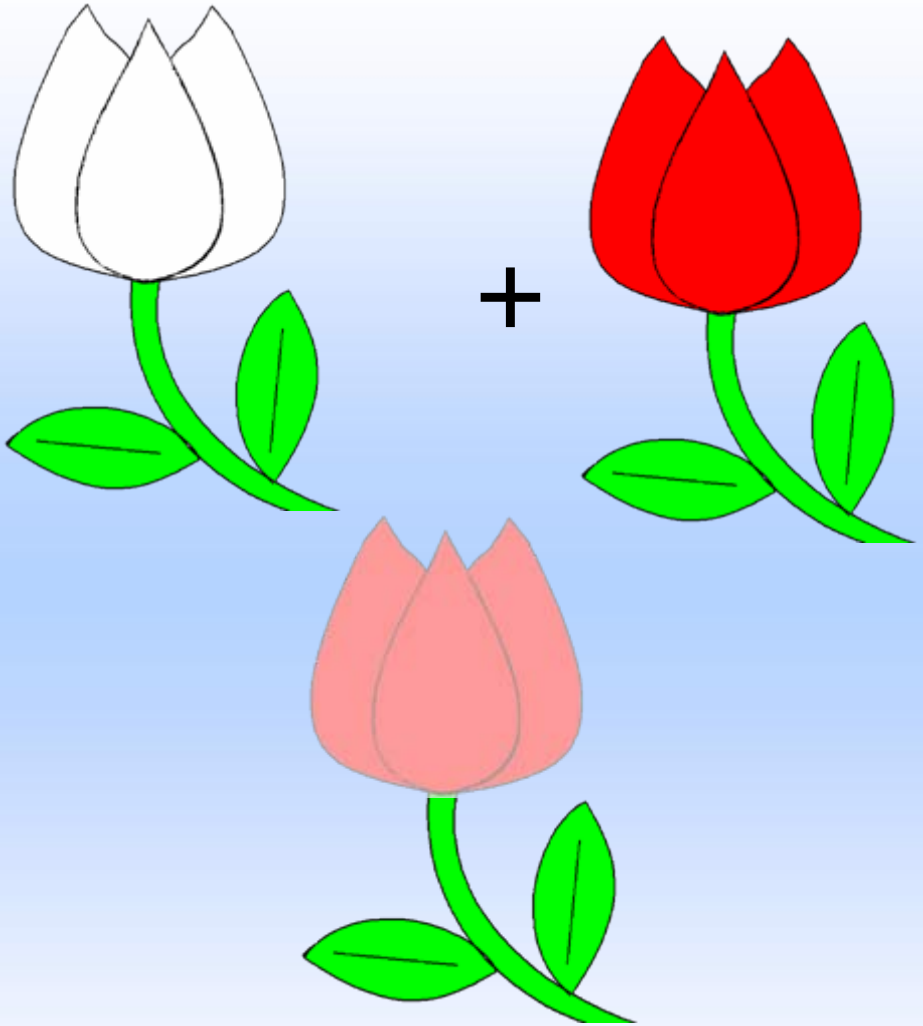
# The Pre-Mendelian Period

- Vital force
- Cells
  - Male heredity


Anton van Leeuwenhoek

# The Pre-Mendelian Period

- Vital force
- Male heredity
- Blending theory



# The Pre-Mendelian Period - Summary

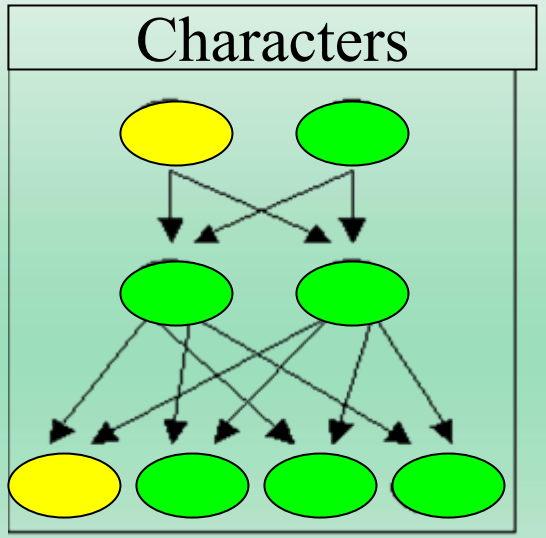
- Middle 19-th century - *Cell Theory*:
  - All living organisms consist of cells
  - Cells of different organisms have similar structure
  - Each new cell is obtained from the parent cell
- Gametes – sexual cells – participate in reproduction, and gametes of both sexes are important.

# An Historical Perspective

- … – 1900 Pre-Mendelian period
- ➢ 1900 – 1940 **Pre-DNA period**
- 1940 – 1990 DNA period
- 1990 – 2003 Genomic period
- 2003 – … Post-genomic era

# The Pre – DNA Period

- 1865 - Gregor Mendel : discrete heredity



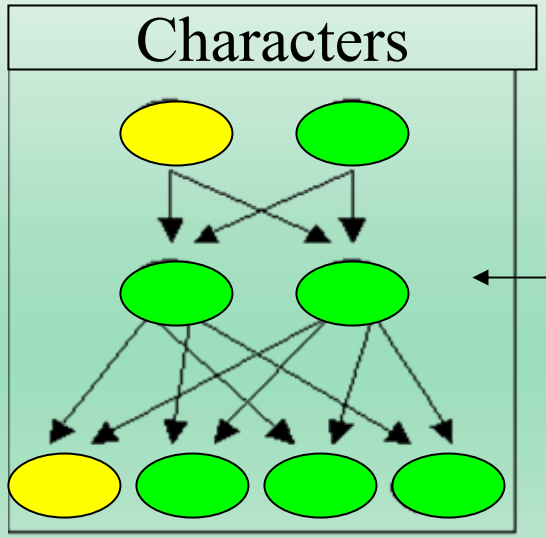

Pure lines

Inbreeding

Appeared again in proportion 3:1

# The Pre – DNA Period

- 1865 - Gregor Mendel : discrete heredity – *elementum - gene*

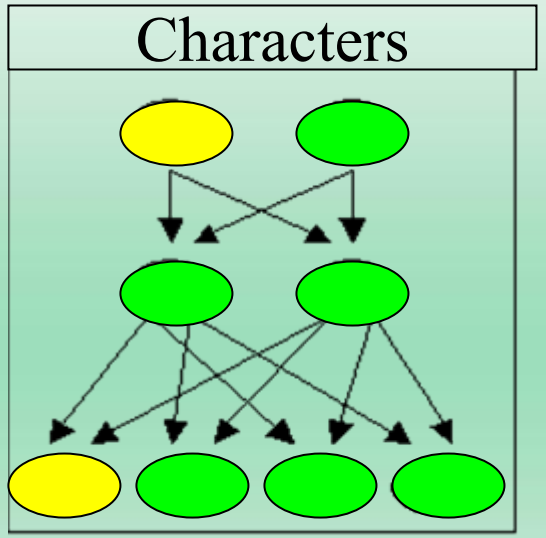



All are *Gy* - green

| Gametes | | *GG*-green | |
|---|---|---|---|
| | | *G* | *G* |
| *yy*- yellow | y | *Gy* | *Gy* |
| | y | *Gy* | *Gy* |

# The Pre – DNA Period

- 1865 - Gregor Mendel : discrete heredity – *elementum - gene*




| Gametes | | *Gy*-green | |
|---|---|---|---|
| | | *G* | *y* |
| *Gy*-green | G | *GG* | *Gy* |
| | y | *Gy* | *yy* |

1/3 is *yy* - yellow

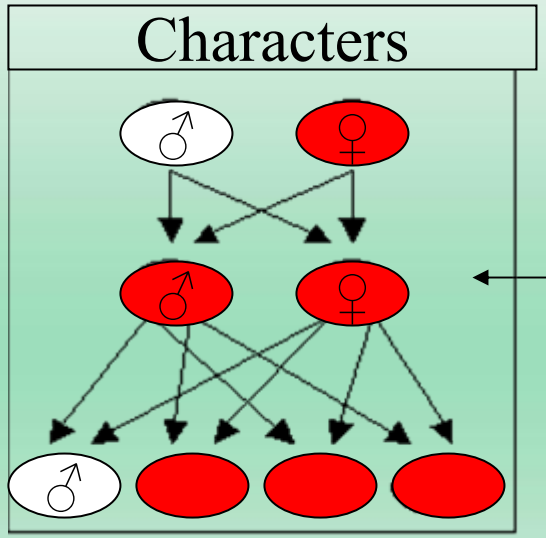
Accepted only in ~1900

# Nucleus and Chromosomes

- Each species contains a specific, and always even, number of chromosomes
- Gametes contain half number of chromosomes.
- The cells of males differ from females by the shape of 1 chromosome. Specifically female carries XX chromosomes, while male carries XY.
- Chromosomes are distributed between two daughter cells in cell division

# The Pre – DNA Period

- 1908 - Thomas Morgan – genes reside on chromosomes



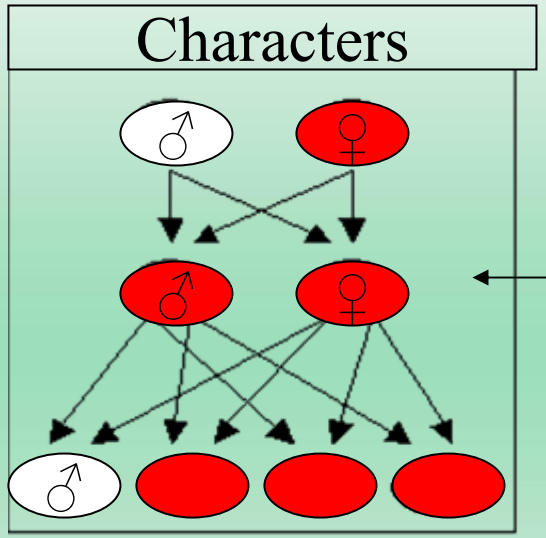

♀: $X^+X^+$ ♂: $X^-Y^-$

All have red eyes

| Gametes | | $X^+X^+$-red | |
|---|---|---|---|
| | | $X^+$ | $X^+$ |
| $X^-Y^-$- white | $X^-$ | $X^+X^-$ | $X^+X^-$ ♀ |
| | $Y^-$ | $X^+Y^-$ | $X^+Y^-$ ♂ |

# The Pre – DNA Period

- 1908 - Thomas Morgan – genes reside on chromosomes




♀: $X^+X^+$ ♂: $X^-Y^-$

cross males with females

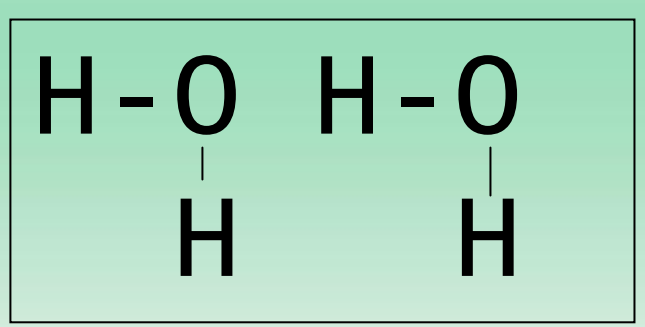
| Gametes | | $X^+X^-$-red | |
|---|---|---|---|
| | | $X^+$ | $X^-$ |
| $X^+Y^-$- red | $X^+$ | $X^+X^+$ | $X^+X^-$ ♀ |
| | $Y^-$ | $X^+Y^-$ | $X^-Y^-$ ♂ |

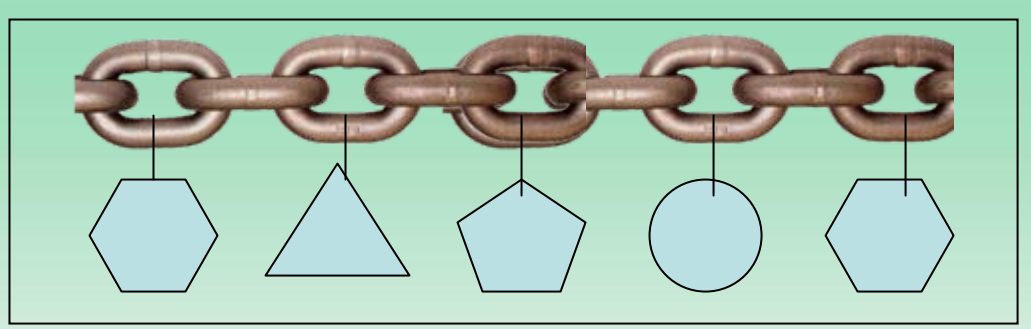
# The Pre – DNA Period

*The geneticist himself is helpless to analyze these properties further. Here the physicist, as well as the chemist, must step in. Who will volunteer to do so?* (Muller 1936, 214)

# The chemical nature of genetic material

- Water?
- Nucleic acid?
- Protein?



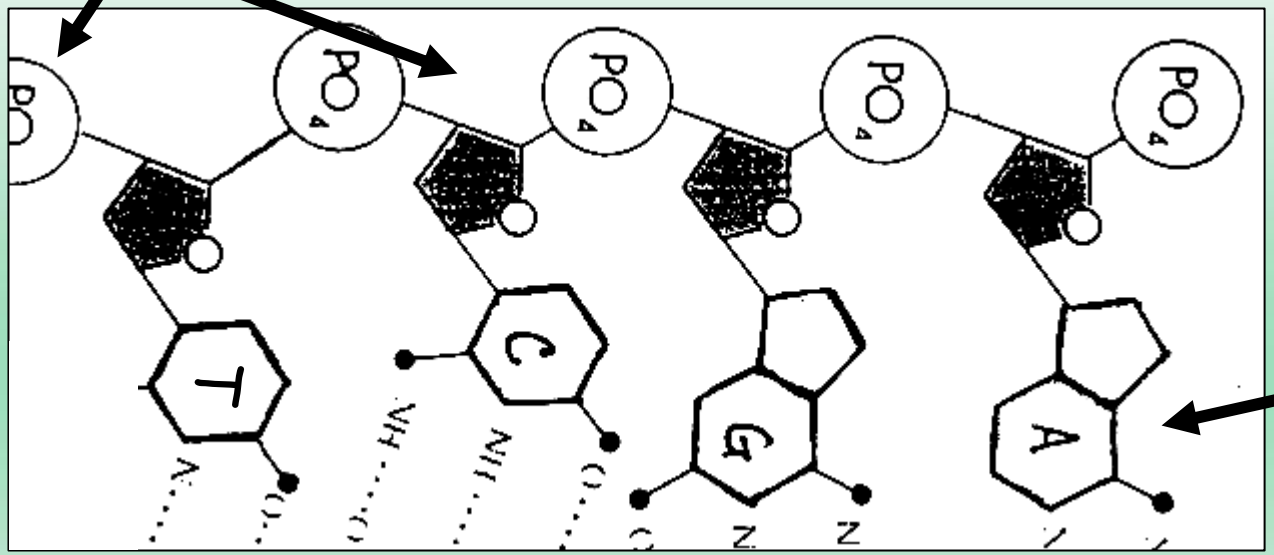



Polymer

# The chemical nature of genetic material

1868 - Friedrich Miescher – *nuclein* - **D**eoxyribo**N**ucleic **A**cid

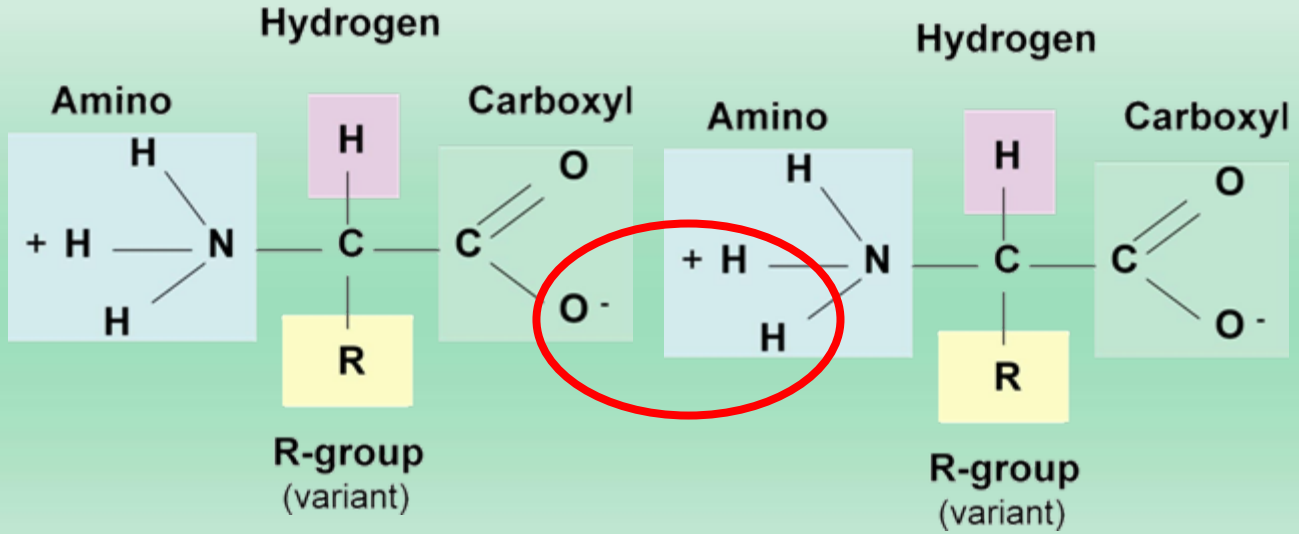
Sugar-phosphate backbone

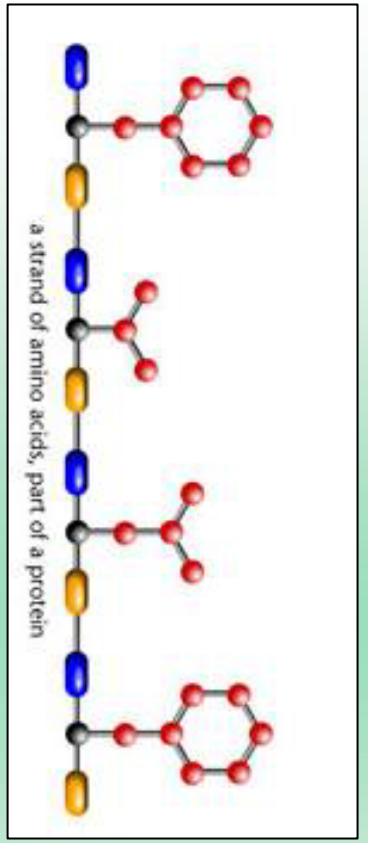


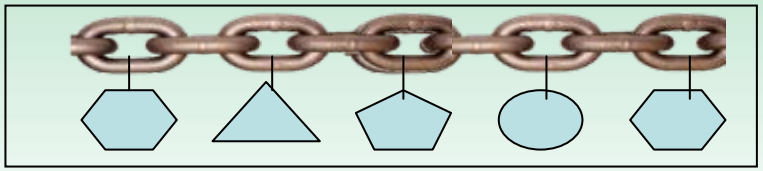

Regular polymer – cannot carry useful information

# The chemical nature of genetic material

Protein – polymer of 20 aminoacids



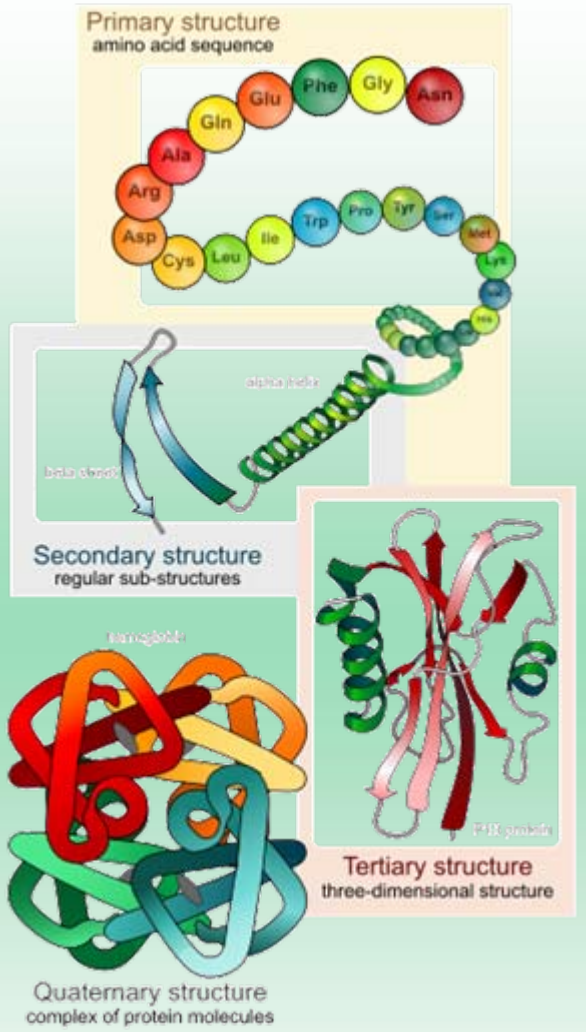



# The chemical nature of genetic material

Proteins come in a variety of functions and shapes (enzymes and structural proteins)

More likely, proteins carry the genetic information



# The Pre – DNA Period

- 1850s – Charles Darwin – *the theory of evolution*
  - Common ancestry
  - Descent with modification

The mechanism of passing traits from generation to generation

# The Pre – DNA Period - Summary

- Genes are the discrete units of heredity
- Genes generate the enzymes that control structural and metabolic functions
- Genes reside on chromosomes
- Two candidates for gene encoding: DNA and protein

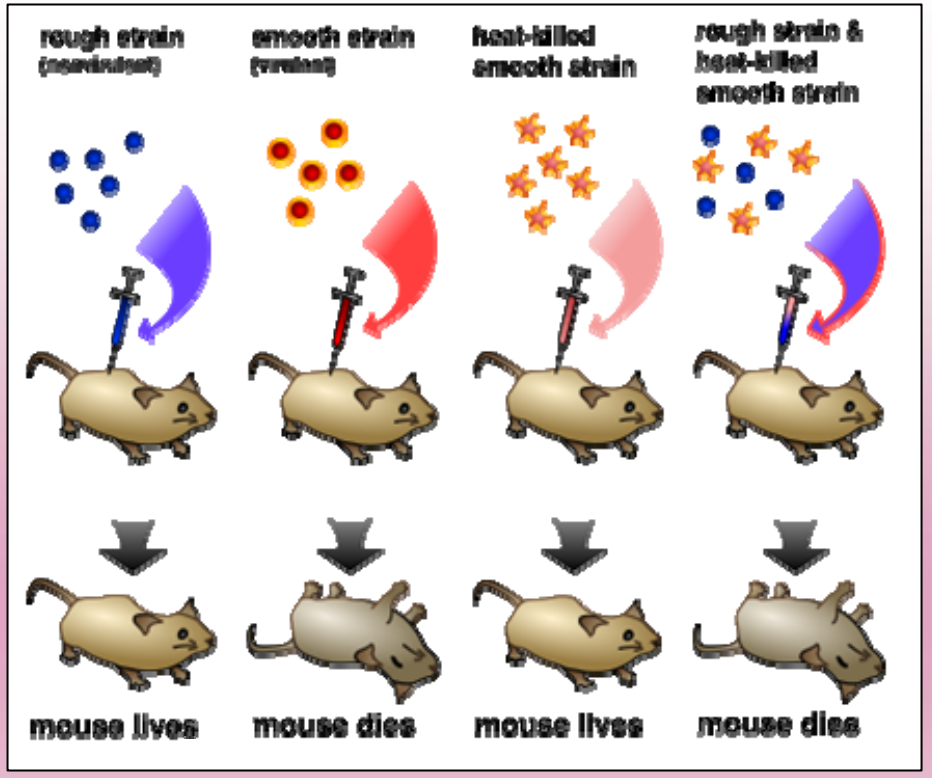
# An Historical Perspective

- … – 1900 Pre-Mendelian period
- 1900 – 1940 Pre-DNA period
- ➢ 1940 – 1990 **DNA period**
- 1990 – 2003 Genomic period
- 2003 – … Post-genomic era

# The DNA Period

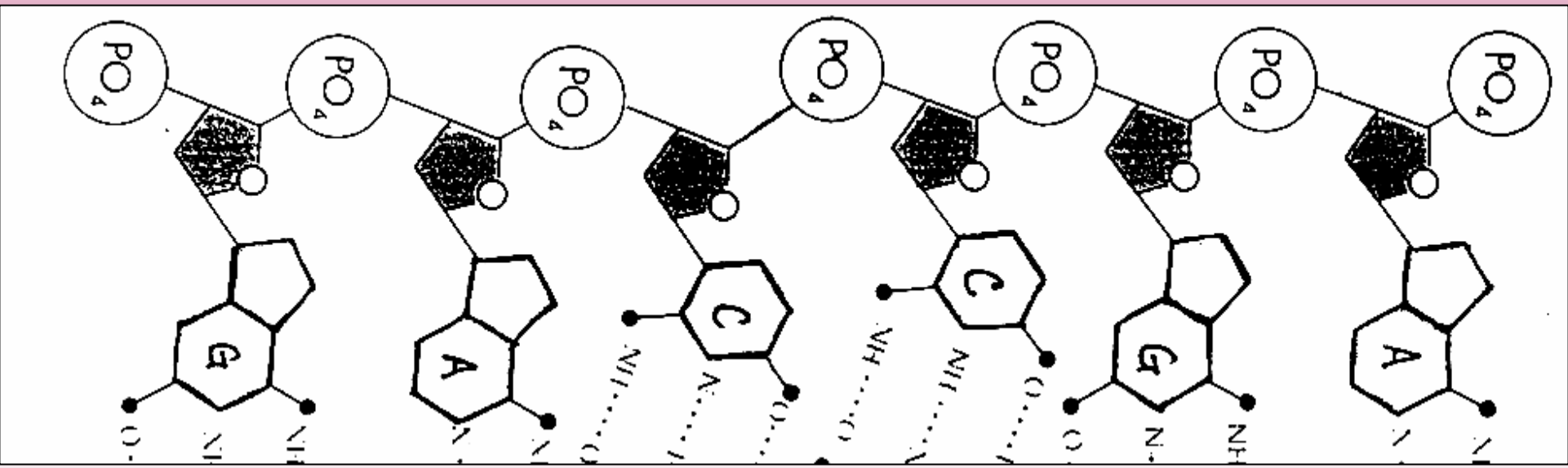
- 1936 – Oswald Avery – experiments with Pneumonia - causing bacteria:
  - rough – harmless
  - smooth - pathogenic

- Explanation: the DNA from the dead pathogenic bacteria used the live non-pathogenic bacteria to synthesize the pneumonia toxins
- The genetic material does not seem to be a protein, since the protein is destroyed by heat



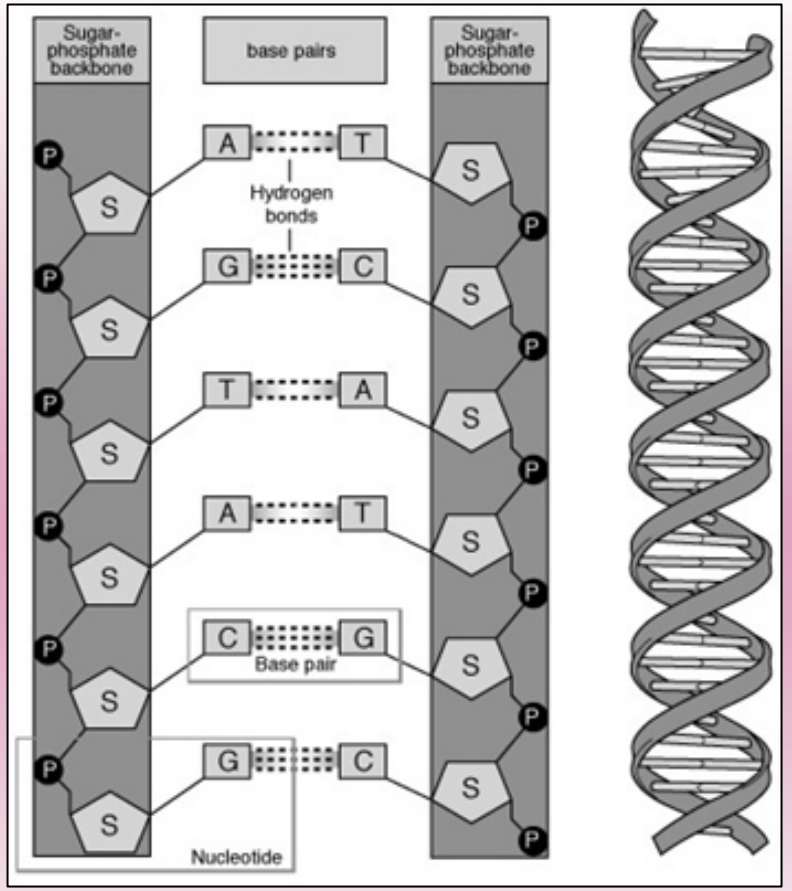
# The DNA Period

- The amount of A = T, C = G
- DNA is an ***irregular*** polymer
- Each species contains the specific amount of DNA
- Viruses – carriers of DNA

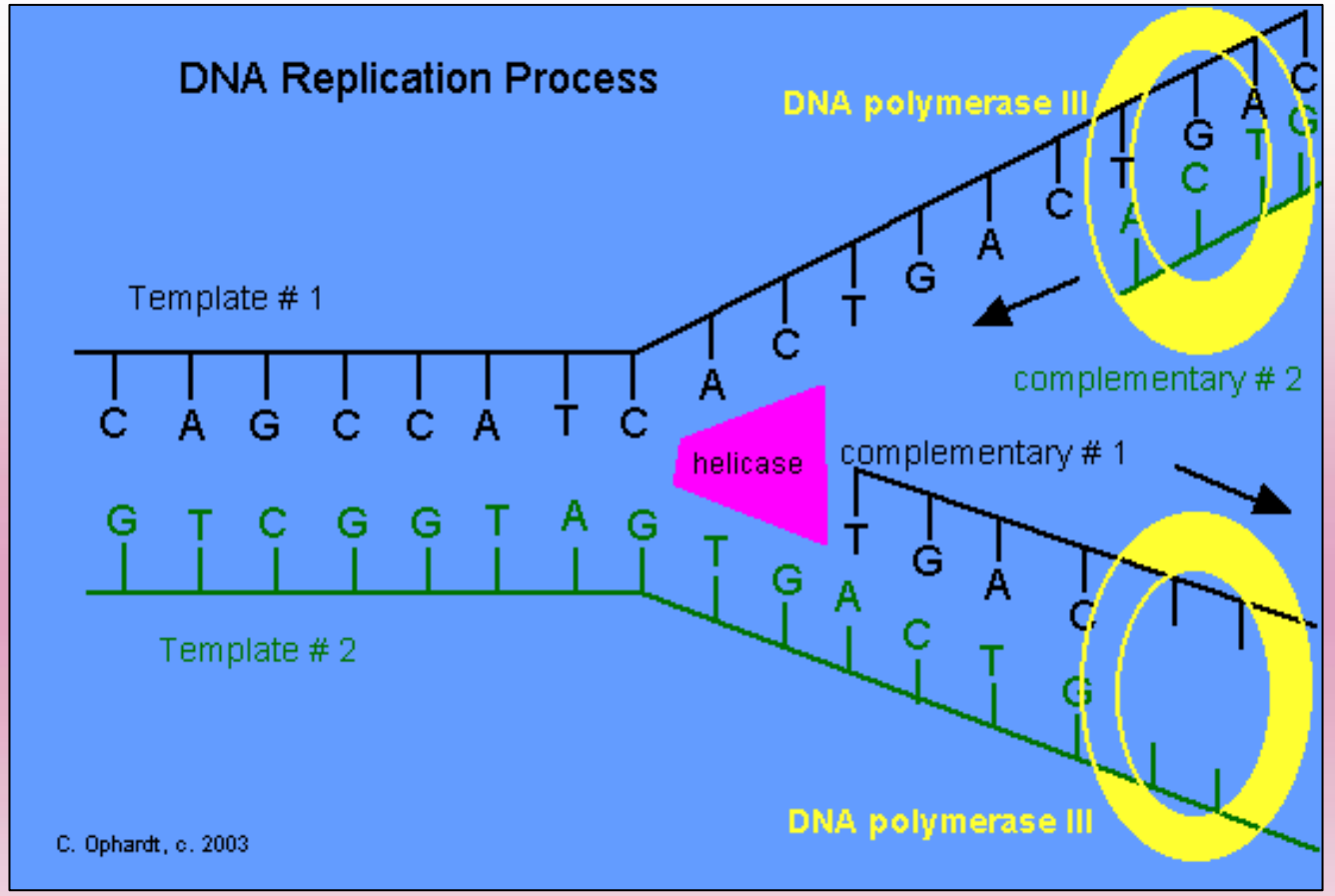


# The DNA Period

- 1953 - James Watson and Francis Crick - double-helix model
- “The structure is good for replication”

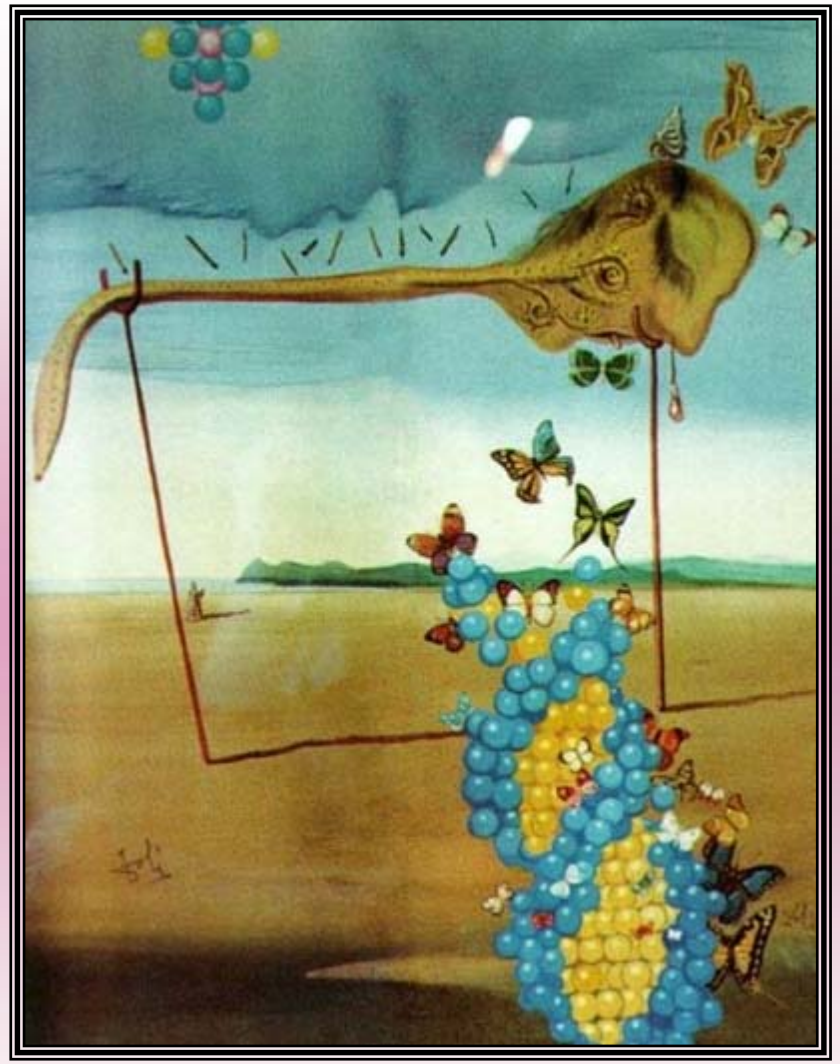


# The DNA Period




Detailed video of DNA replication

# DNA – a secular icon of modern society

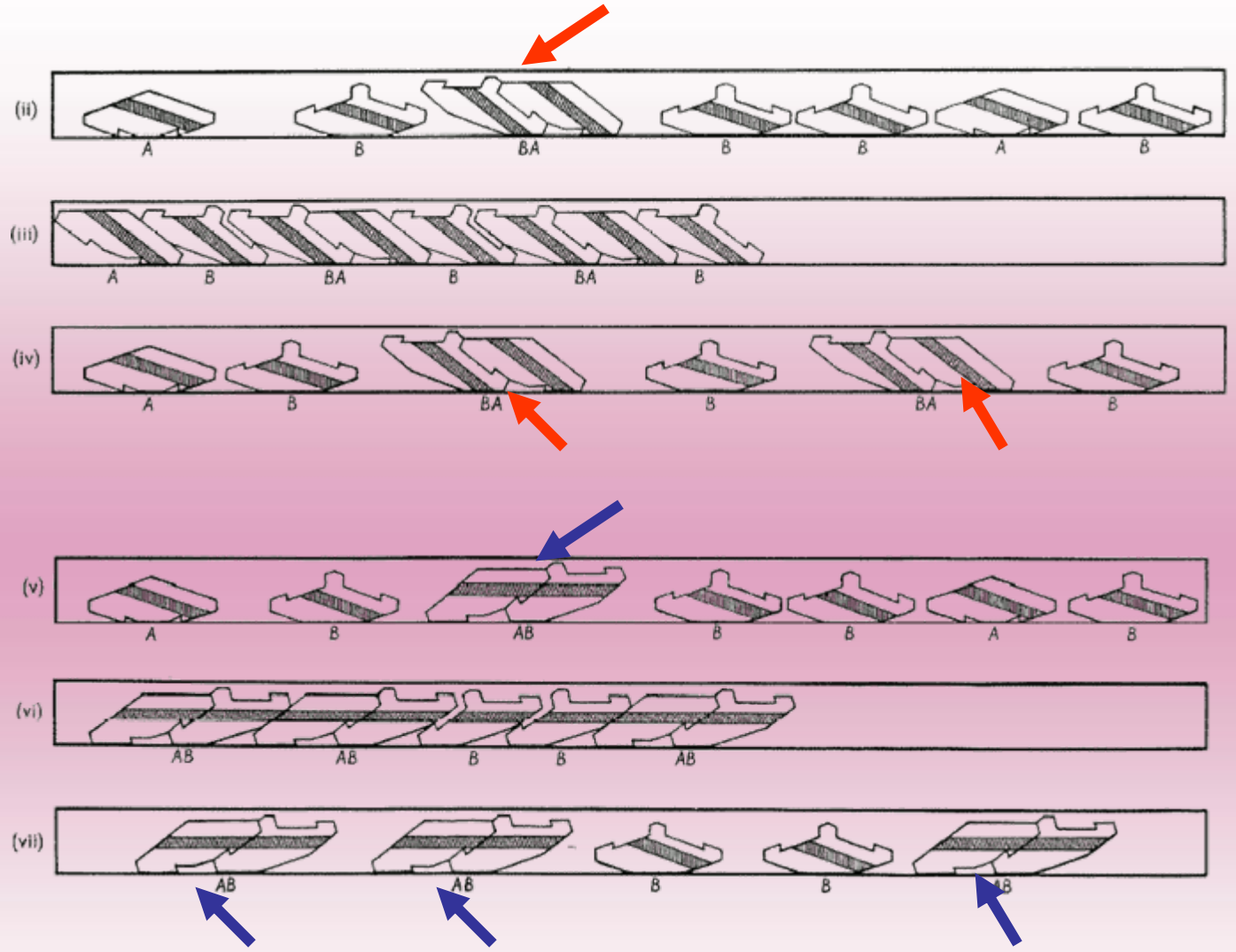


Salvatore Dali

*Butterfly Landscape*

created only a few years after Watson and Crick's announcement of the double-helix

# Mechanical Self-Replicating Systems



# Mechanical Self-Replicating Systems

*It might have been thought that, to achieve mechanical self-replication, some form of magnetic or electrical forces could have been conveniently used. This was a point carefully considered and it seemed in fact much simpler to use gravity, friction and collision as the forces determining the hooking and activating mechanisms which were evolved.*

L.S. Penrose. Mechanisms of self-reproduction. 1958

**Link to Penrose's paper**

# DNA replication

- Semi-conservative
- Errors-prone: substitutions

| | A | T | C | G |
|---|---|---|---|---|
| **A** | 0 | 5 | 5 | 1 |
| **T** | 5 | 0 | 1 | 5 |
| **C** | 5 | 1 | 0 | 5 |
| **G** | 1 | 5 | 5 | 0 |

A, G – 2-ring bases

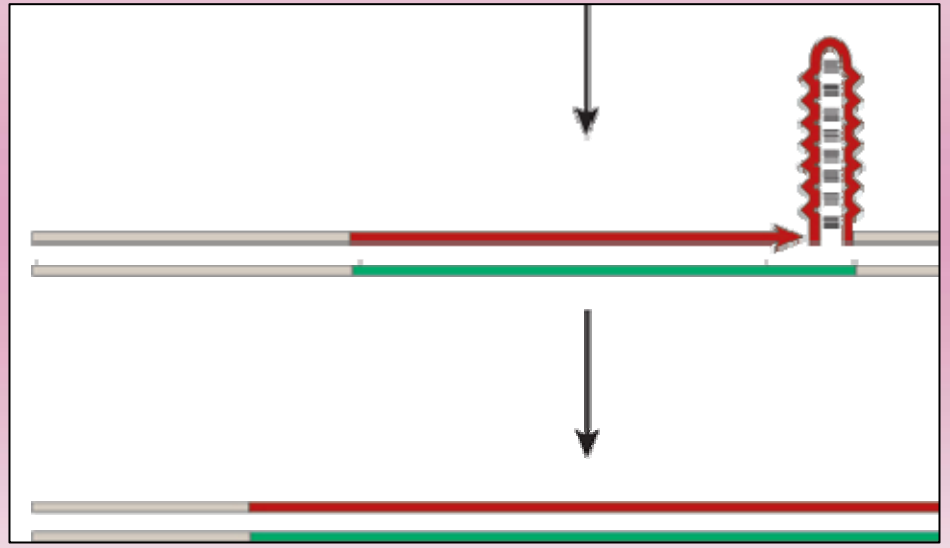
T,C – 1-ring bases

Mutation which preserves rings number is much more likely than changing the rings number.

The score of using G instead of A is (1) - less than the score of using T (5) instead of A

# DNA replication

- Semi-conservative
- Errors-prone: deletions



Hairpin loop

# DNA replication

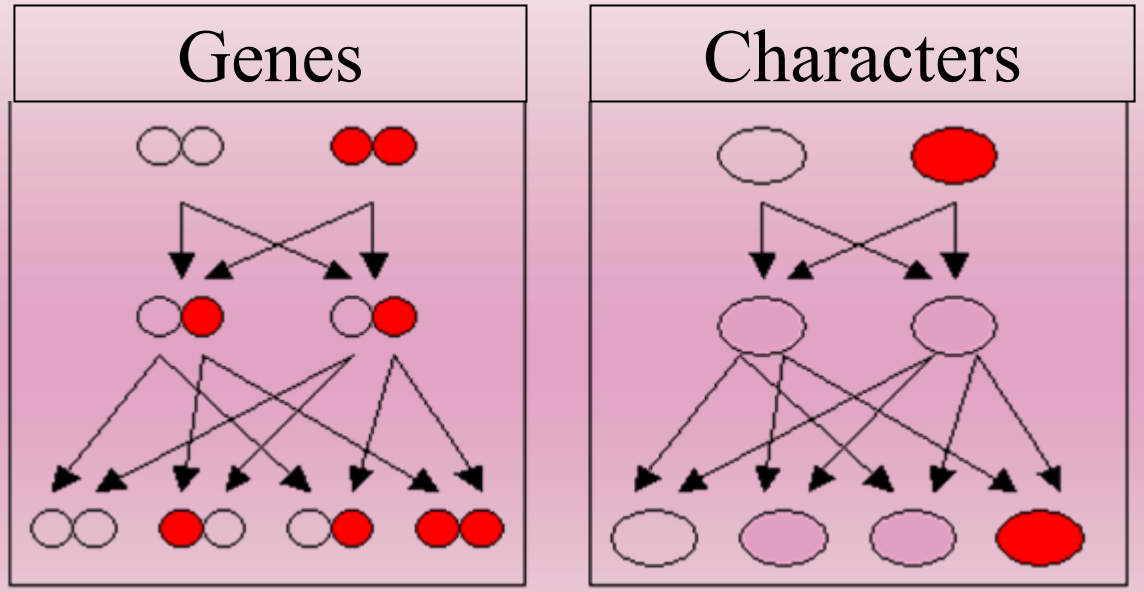
- Semi-conservative
- Errors-prone: point mutations, deletions, breaks
- DNA repair mechanisms

- This leads to the relative
  - *stability* of the DNA molecule over generations
  - and the ability to *change*

the basic mechanisms of evolution

# From gene to protein

- *Phenotype* – an outward expression of discrete genetic characteristics. Proteins are responsible for phenotype




How information from the sequence of nucleotides is converted into a sequence of aminoacids ?

# Protein Synthesis: Transcription

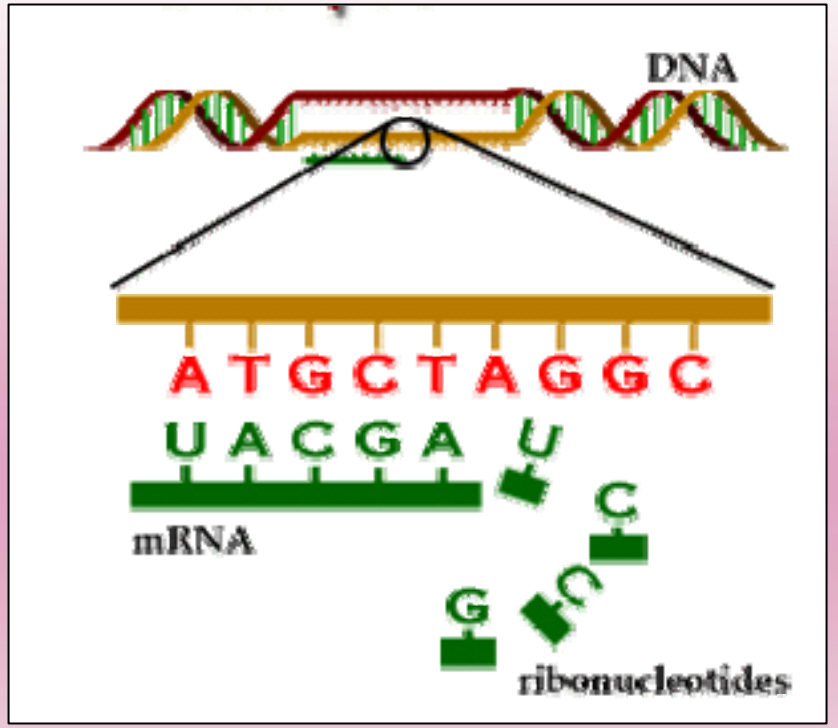
- RNA – **R**ibo**N**ucleic **A**cid -

  a short and unstable polymer of the same nucleotides as DNA:

  **A**denine, **C**ytosine, **G**uanine, **U**racil (instead of **T**hymine)

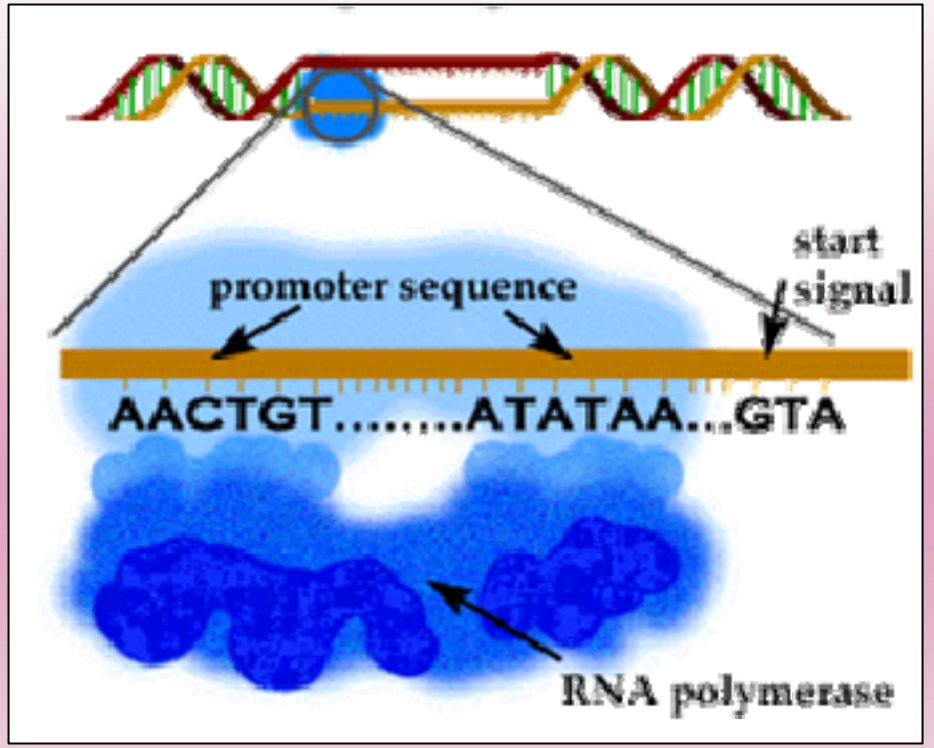
Messenger-RNA, **m-RNA**

- Copy of the template strand of DNA is made in the cell nucleus
- The copy moves into cytoplasm



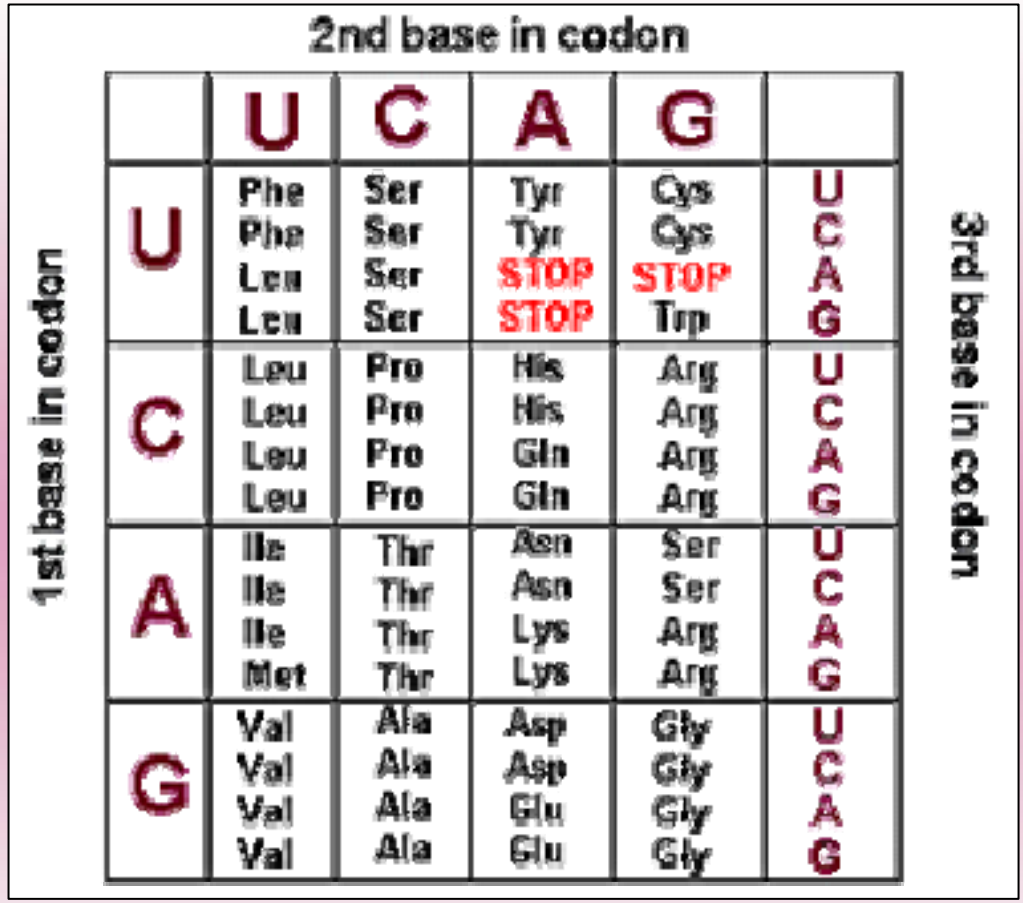
# Protein Synthesis: Transcription

- Initiation with binding of the RNA polymerase to the ***promoter*** site (comparatively conserved sequences).
- The synthesis starts at ***start codon*** GTA (which then become bases CAU on the RNA molecule).



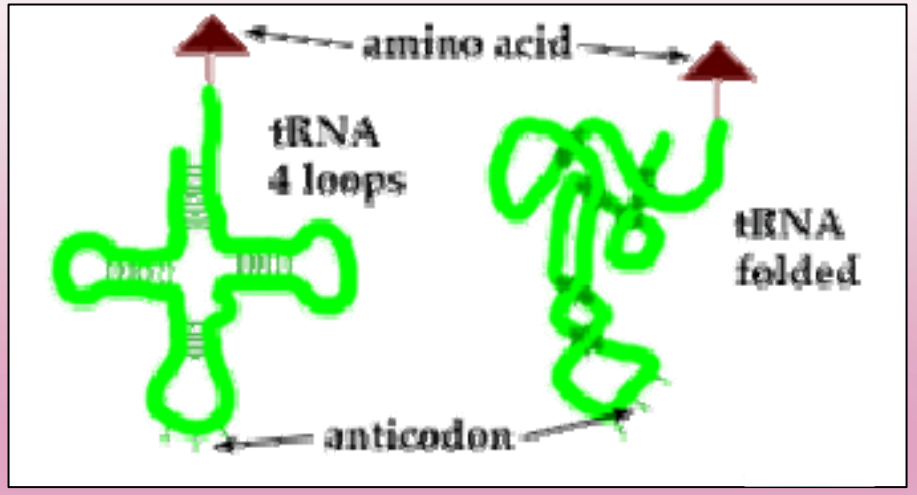
# Genetic Code

- There are $4^3$=64 possible triplets – *codons*, but only 20 aminoacids and 3 stop codons.
- The code is degenerative: different triplets code for the same aminoacid
- Important in keeping the proteins functional



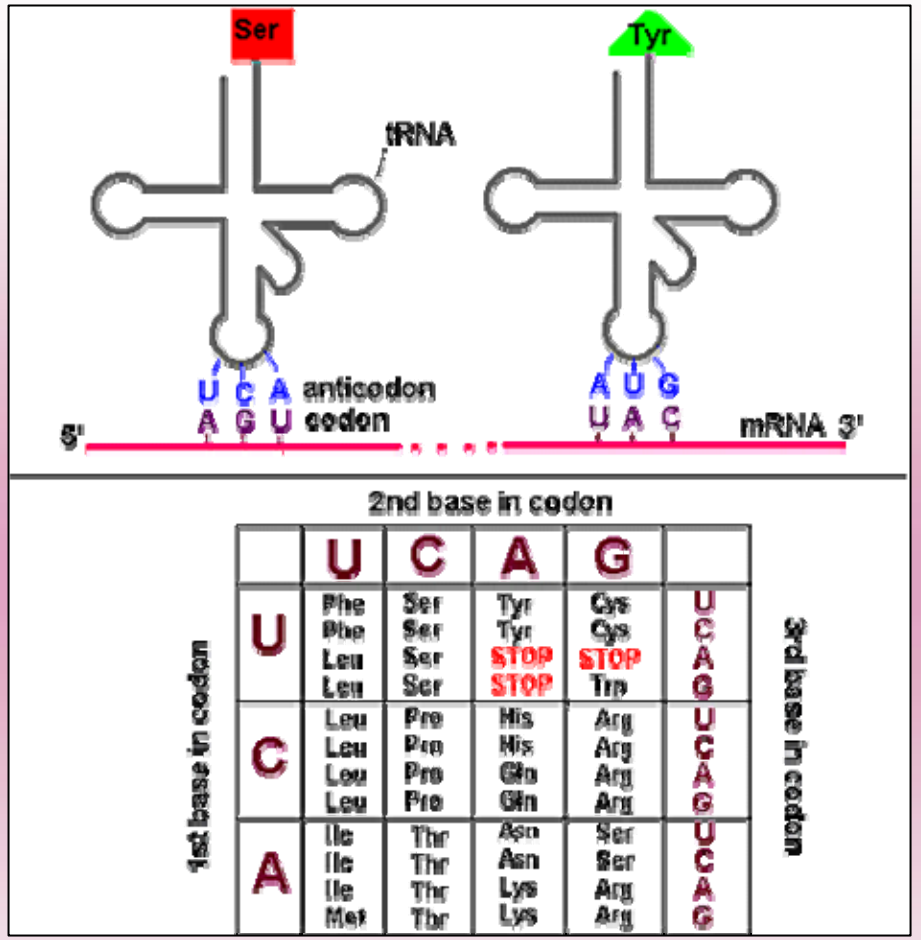
# Protein Synthesis: Transport RNA

- **t-RNAs** are short
- Fold into a cloverleaf secondary structure
- Hydrogen bonds hold into an L-shaped tertiary structure

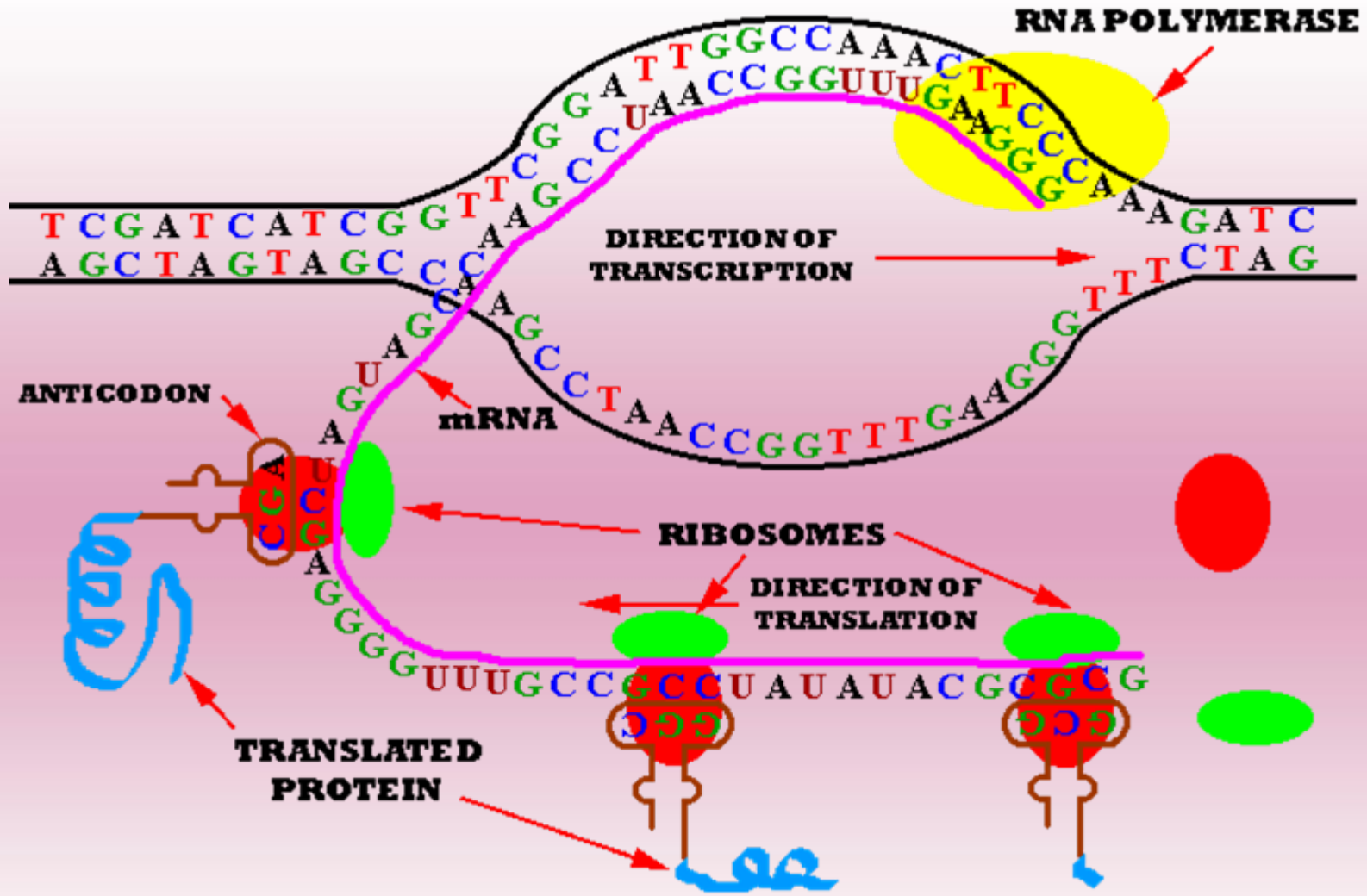


# Protein Synthesis: Transport RNA

The anticodon is complementary to the triplet encoding the attached aminoacid, according to the genetic code



# Protein Synthesis



# The central dogma of molecular biology

DNA contains the complete genetic information that defines the structure and function of an organism. Proteins are formed using the genetic code of the DNA. Three different processes are responsible for the inheritance of genetic information and for its conversion from one form to another :

1. Replication
2. Transcription
3. Translation

# The central dogma of molecular biology

| General | Special | Unknown |
|---|---|---|
| DNA → DNA | RNA → DNA | protein → DNA |
| DNA → RNA | RNA → RNA | protein → RNA |
| RNA → protein | DNA → protein | protein → protein |

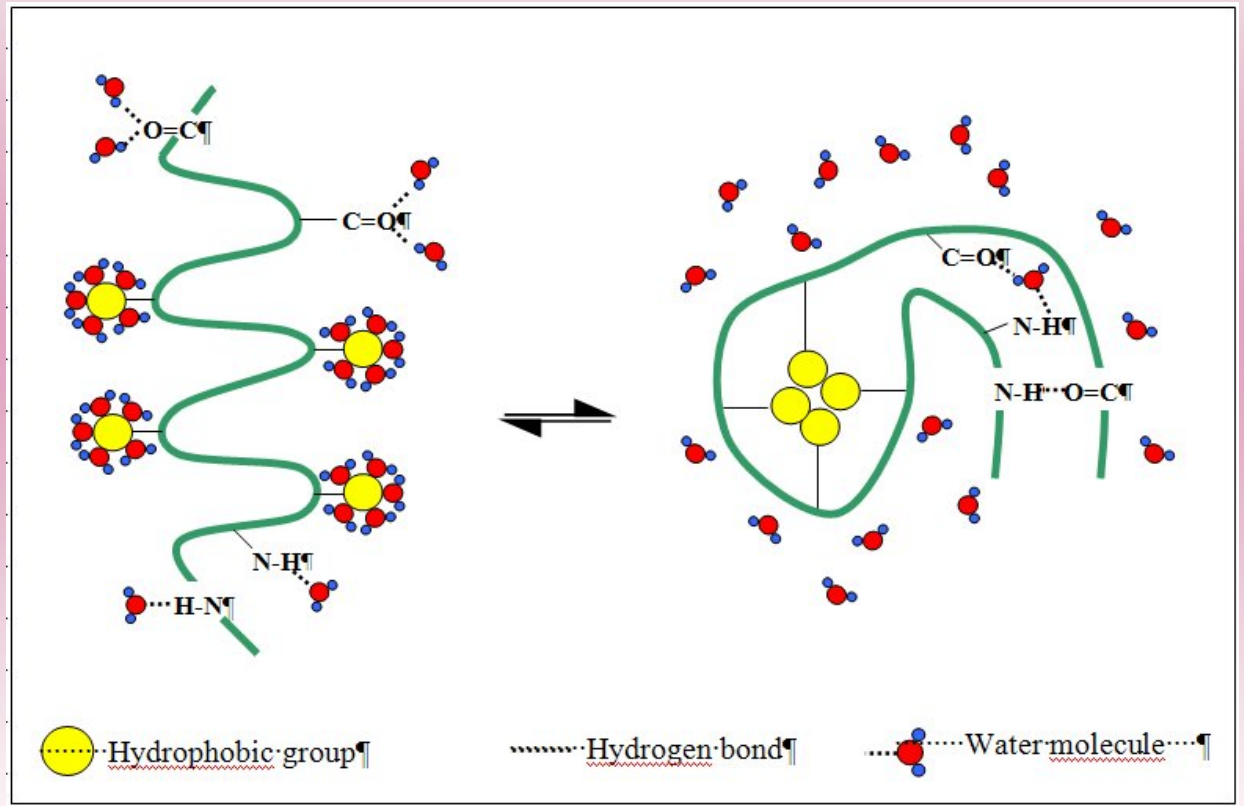
The direction of the information flow:

DNA->RNA->Protein,

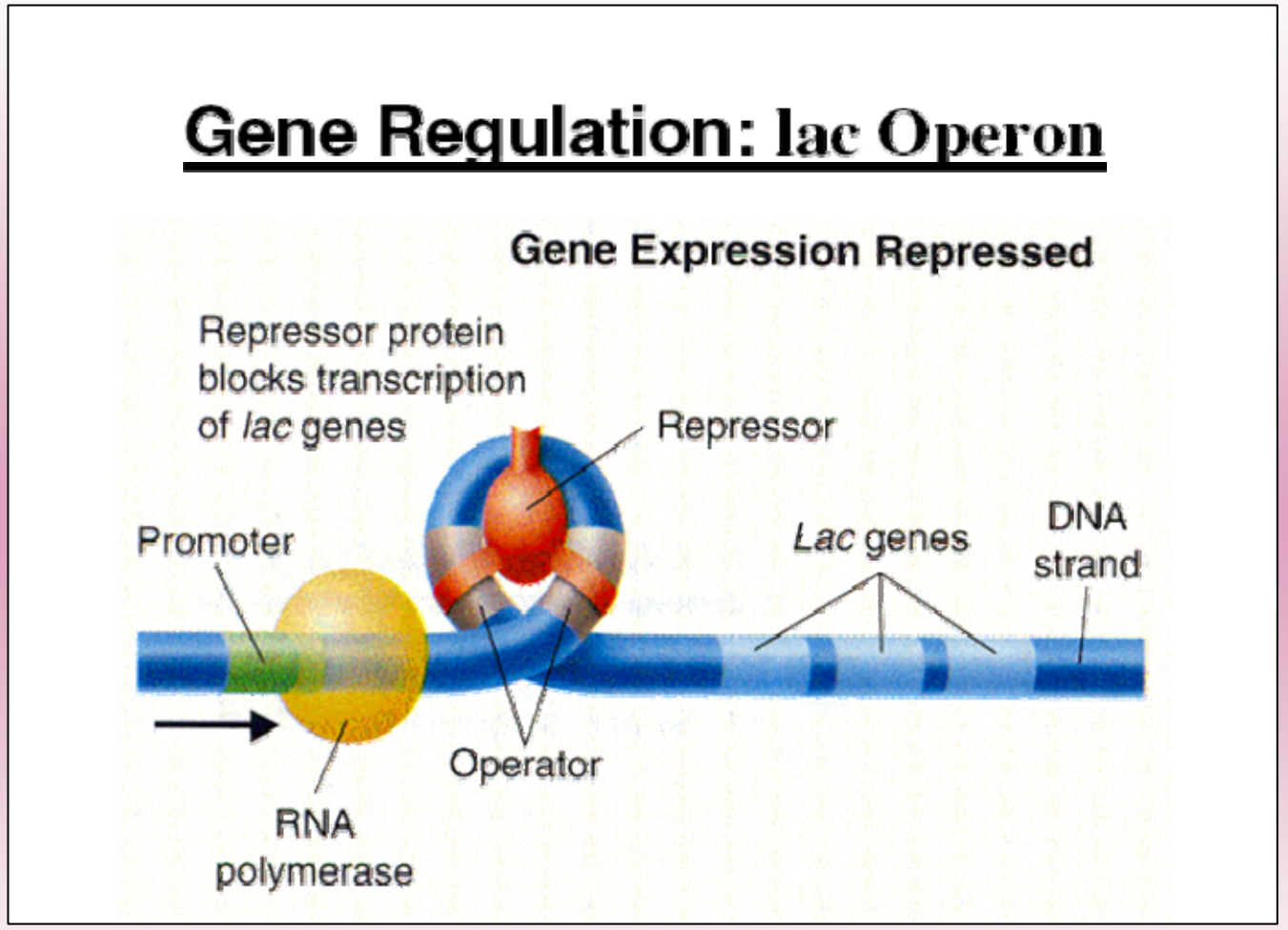
never Protein->DNA

# Protein Folding

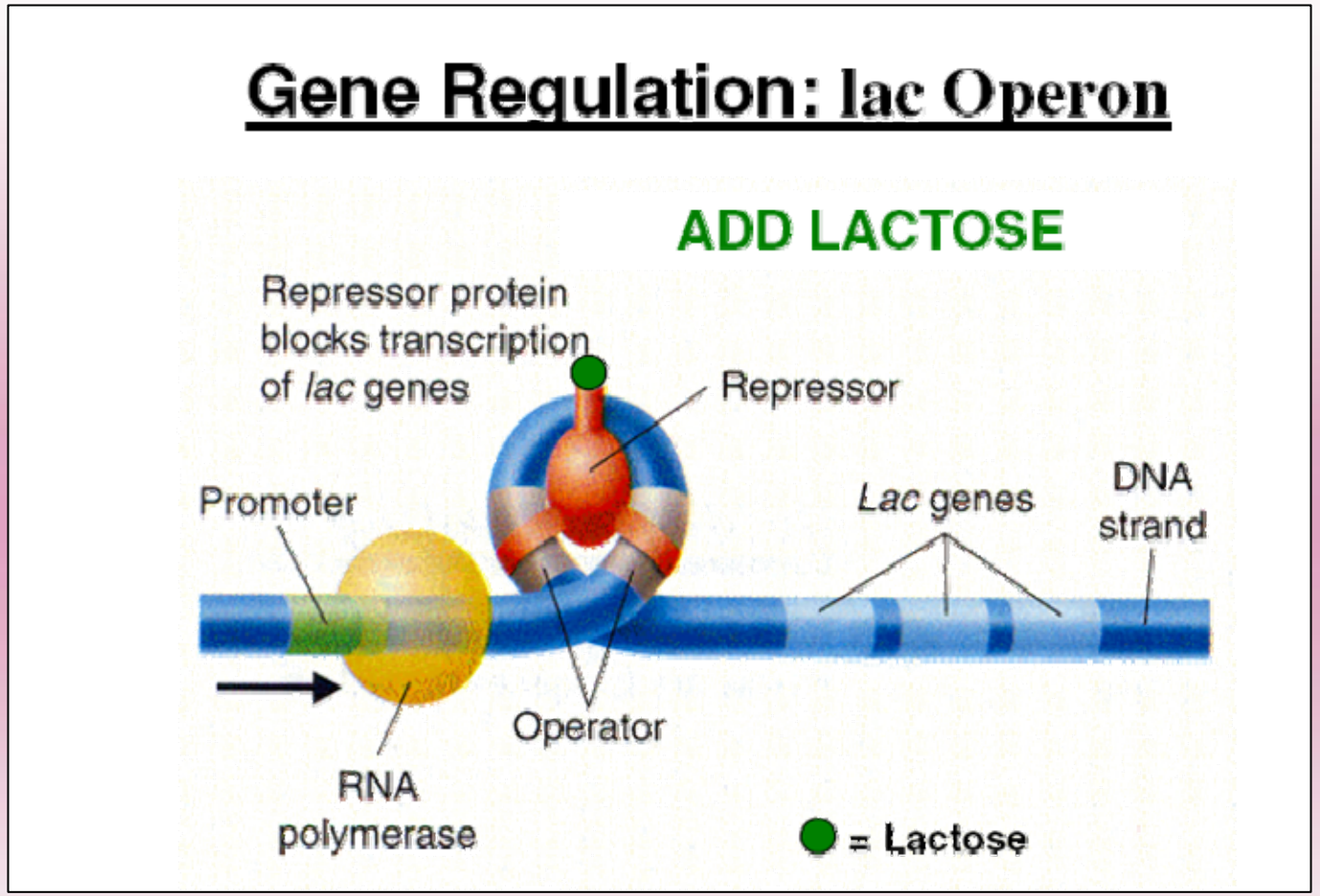
- The property of folding is spontaneous and is determined by the sequence of aminoacids
- The folding is mostly caused by hydrophobic-hydrophilic properties of aminoacid residues, which determine the orientation of these residues in a water environment, plus additional week bonds



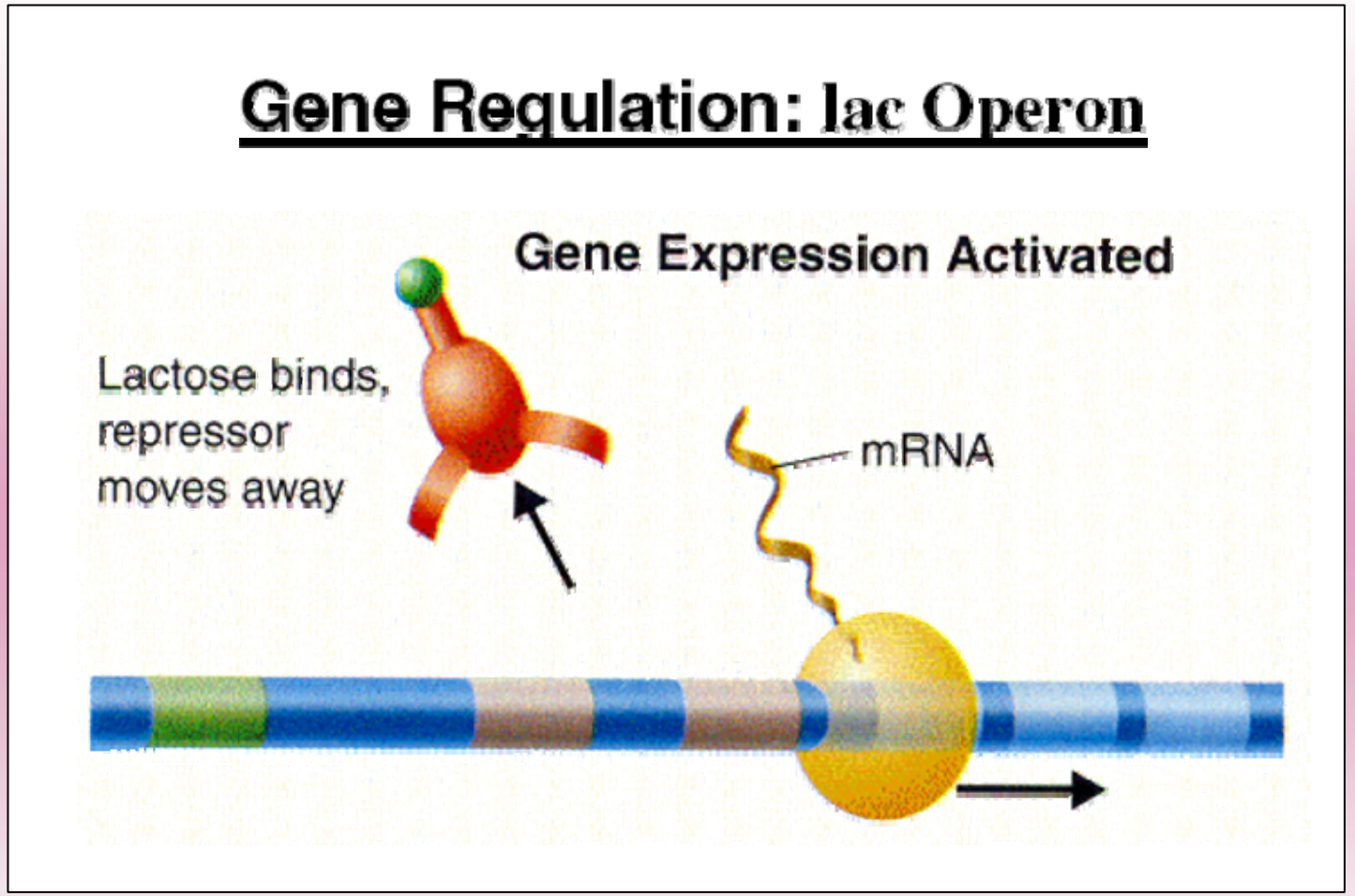
# Regulation of gene expression



# Regulation of gene expression

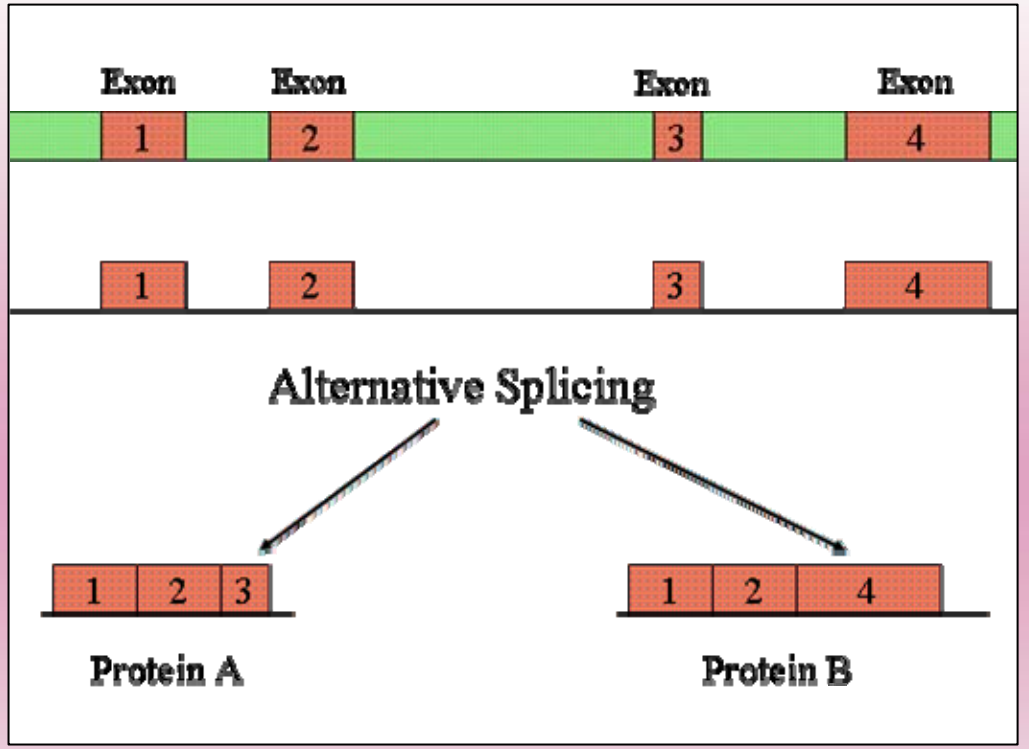


# Regulation of gene expression



# Gene to Protein: Complications

- Collinearity between the linear order of nucleotides and the linear order of aminoacids – did not persist after the 70-s:
  - Overlapping genes – different proteins from the same overlapping sequence of DNA
  - Interleaving *exons* (coding) and *introns* (non-coding) regions
  - Alternative splicing of exons



# Essential Knowledge

- DNA replication – semi-conservative, mutations, repair, stability and change
- Transcription from DNA to m-RNA – volatile RNA from stable DNA
- Splicing of m-RNA transcript into protein-coding sequence
- Translation – t-RNA, genetic code, degenerative code, stop-codons, promoters
- The central dogma of molecular biology – direction of the information flow

# The DNA Period - Summary

- DNA – a polymer of 4 types of nucleotides – carries all the information needed for life
- This information is expressed through synthesis of proteins. The aminoacid sequence of each protein determines its shape, which determines the function of the protein.