Bioinformatics

CS300

Crash course:

Structure and Replication of DNA

Spring 2021

Oliver BONHAM-CARTER

An Informatics by Another Name?

- **Bio-medical Informatics**
  - Study population-level data to manage health
- **Clinical informatics**
  - Systems used to deal with patient health
  - Clinical trial management systems, electronic health records, etc.
- **Laboratory information**
  - Systems to deal with scientific instrumentation and data management
  - Connecting instruments together, managing laboratory flow, etc.
- **Bioinformatics**
  - Systems to study DNA to answer questions about disease, relatedness, health and other concepts
  - DNA, RNA, proteins, *molecular* systems by data

A Current Need For Bioinformatics?

- Vaccine research: reduce the time for developement

  "The Impact of Bioinformatics on Vaccine Design and Development", by Ribas-Aparicio *et al.* *link*

Abstract

Vaccines are the pharmaceutical products that offer the best cost-benefit ratio in the prevention or treatment of diseases. In that a vaccine is a pharmaceutical product, vaccine development and production are costly and it takes years for this to be accomplished. Several approaches have been applied to reduce the times and costs of vaccine development, mainly focusing on the selection of appropriate antigens or antigenic structures, carriers, and adjuvants. One of these approaches is the incorporation of bioinformatics methods and analyses into vaccine development. This chapter provides an overview of the application of bioinformatics strategies in vaccine design and development, supplying some successful examples of vaccines in which bioinformatics has furnished a cutting edge in their development. Reverse vaccinology, immunoinformatics, and structural vaccinology are described and addressed in the design and development of specific vaccines against infectious diseases caused by bacteria, viruses, and parasites. These include some emerging or re-emerging infectious diseases, as well as therapeutic vaccines to fight cancer, allergies, and substance abuse, which have been facilitated and improved by using bioinformatics tools or which are under development based on bioinformatics strategies.

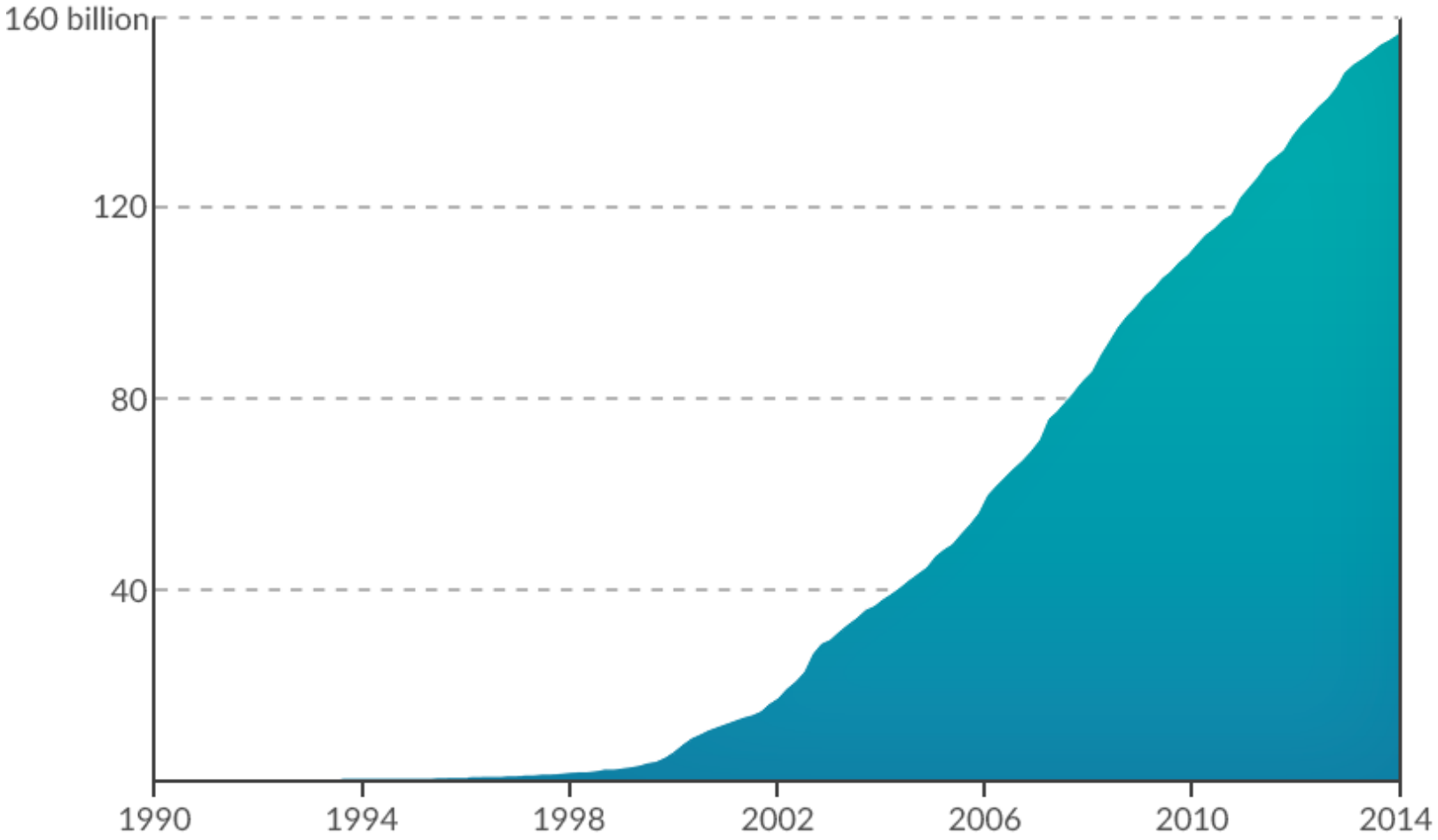
Keywords: reverse vaccinology, immunoinformatics, structural vaccinology, computational strategies, vaccine

Other Needs for Bioinfomatics?

- Avoidance: Better decisions after information of high risks of types of disorders
  - Diabetes, cancer, other disorders
- Research: To process the health data that *already exists*.
  - Open health article (towardsdatascience.com) link
  - What can we learn from our experiments? Past experiments? Can this data train future models?
  - For instance, could we make computer models from data of animal models? (no more animal models?)

Lots of Data to Process

- Exponential Growth of NIH base pairs through December 2013



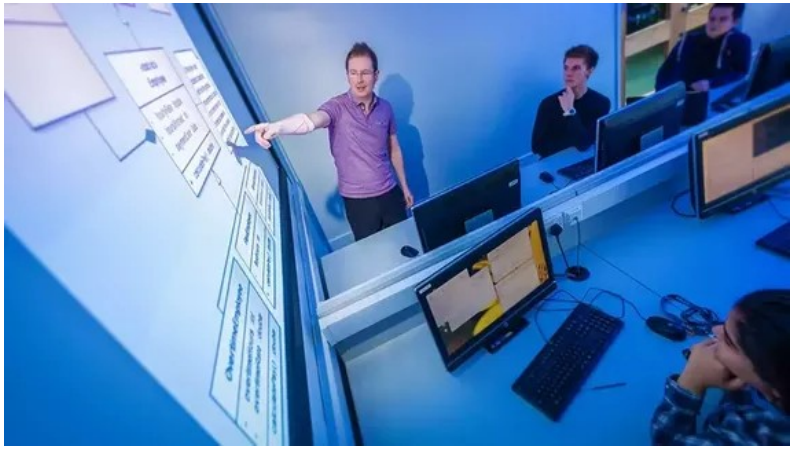
Needed for Careers?

- Biologists will need
  - Programming skills
  - Mathematical /statistical skill
  - Programming for Automation
  - Skills to manage data
  - Others

- Computer Scientists will need
  - Biology knowledge
  - Knowledge of biological systems and mechanisms
  - Skills to build models after biological systems
  - Others

Career Ideas

- Professional search link
- Software (bioinformatics) engineer
- Research scientist in biotechnology
- Data scientist
- Project manager (pharmaceuticals, medical, etc)
- Computational immunologist
- Medical doctor (in clinical and research applications)



Researchers and Scientists

Bioinformatics Scientists are generally graduate-educated research scientists whose work involves the development of computer and technology-aided solutions to problems in biochemistry and biological research.

Primary focus areas may include genomics and proteomics.These professionals may be required to create, maintain and utilize databases of complex biodata,and utilize exsiting publicly available databases containing similar information.

Go online to read more about careers!
https://www.recruiter.com/salaries/bioinformatics-scientists-salary/

Job Listings

28 Feb 2021

ALLEGHENY COLLEGE

glassdoor | Jobs | Company Reviews | Salaries | Interviews | Salary Calculator | Sign In | Write Review | For Employ

Bioinformatics Scientist | Jobs | Location

All Job Types | Posted Any Time | $11K-$217K | 25 Miles | More

Most Relevant — 1365 Bioinformatics Scientist Jobs

GRAIL
Bioinformatics Scientist
Menlo Park, CA
$99K - $160K (Glassdoor est.) — Hiring Surge — 11d
4.0 ★

NeoGenomics Laboratories
Bioinformatics Scientist
Fort Myers, FL
$65K - $107K (Glassdoor est.) — Actively Hiring — 11d
4.1 ★

GSK
Clinical Bioinformatics Investigator (Oncology)
Waltham, MA
$52K - $95K (Glassdoor est.) — New — 3d
4.0 ★

New England Biolabs
Research Scientist I, RNA Division
Ipswich, MA
$51K - $97K (Glassdoor est.) — Hot — 4d
4.9 ★

GRAIL 4.0 ★
Bioinformatics Scientist
Menlo Park, CA
$99K - $160K (Glassdoor est.)

Apply Now

Hiring Surge — This company is in a hiring surge in response to COVID-19

Rating Highlights

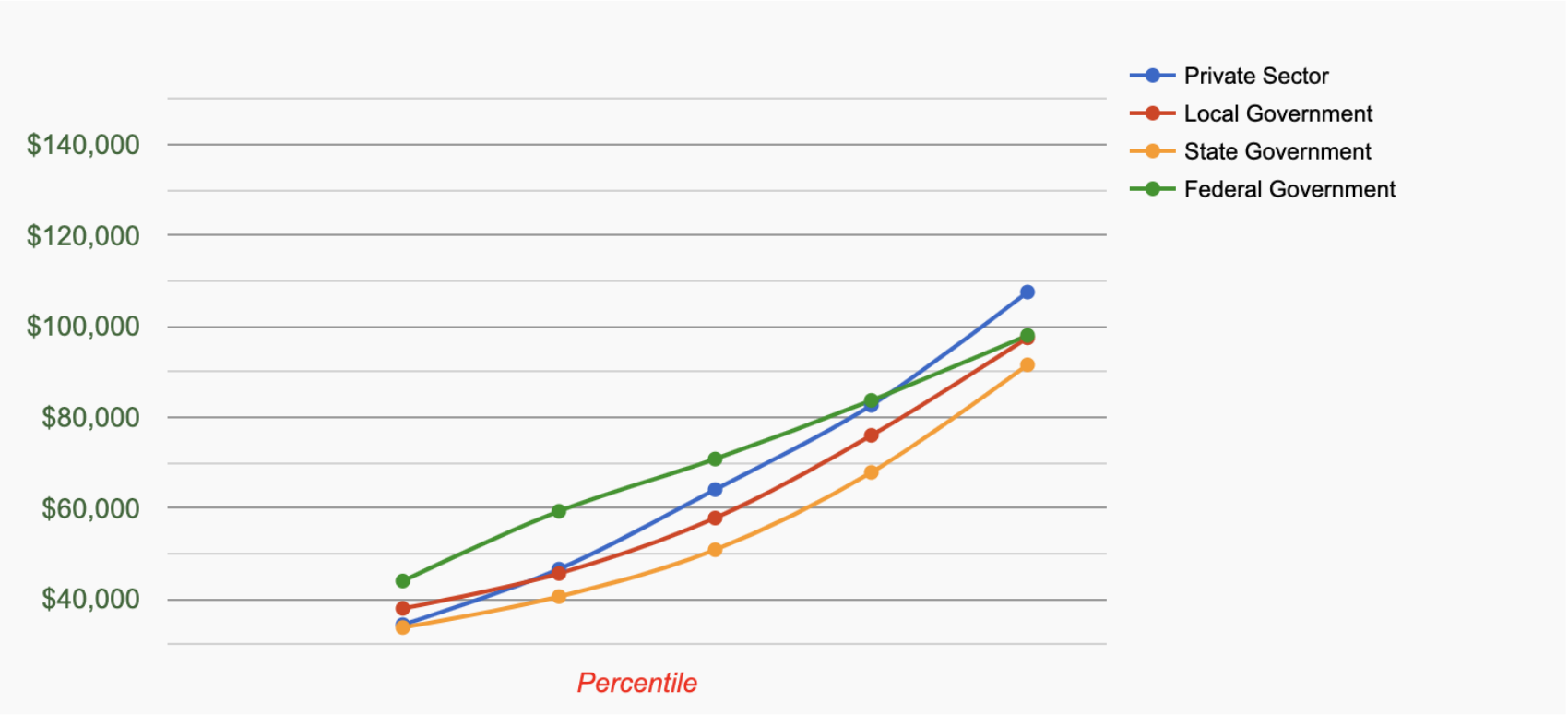
Compensation & Benefits: 3.6 ★
Culture & Values: 3.8 ★
Career Opportunities: 3.9 ★
Work/Life Balance: 3.4 ★

Job & Company Insights

Job Type: Full-time
Job Function: Bioinformatics Sc
Industry: Biotech & Pharmaceu
Size: 201 to 500 Employees

Wages

ANNUAL SALARY



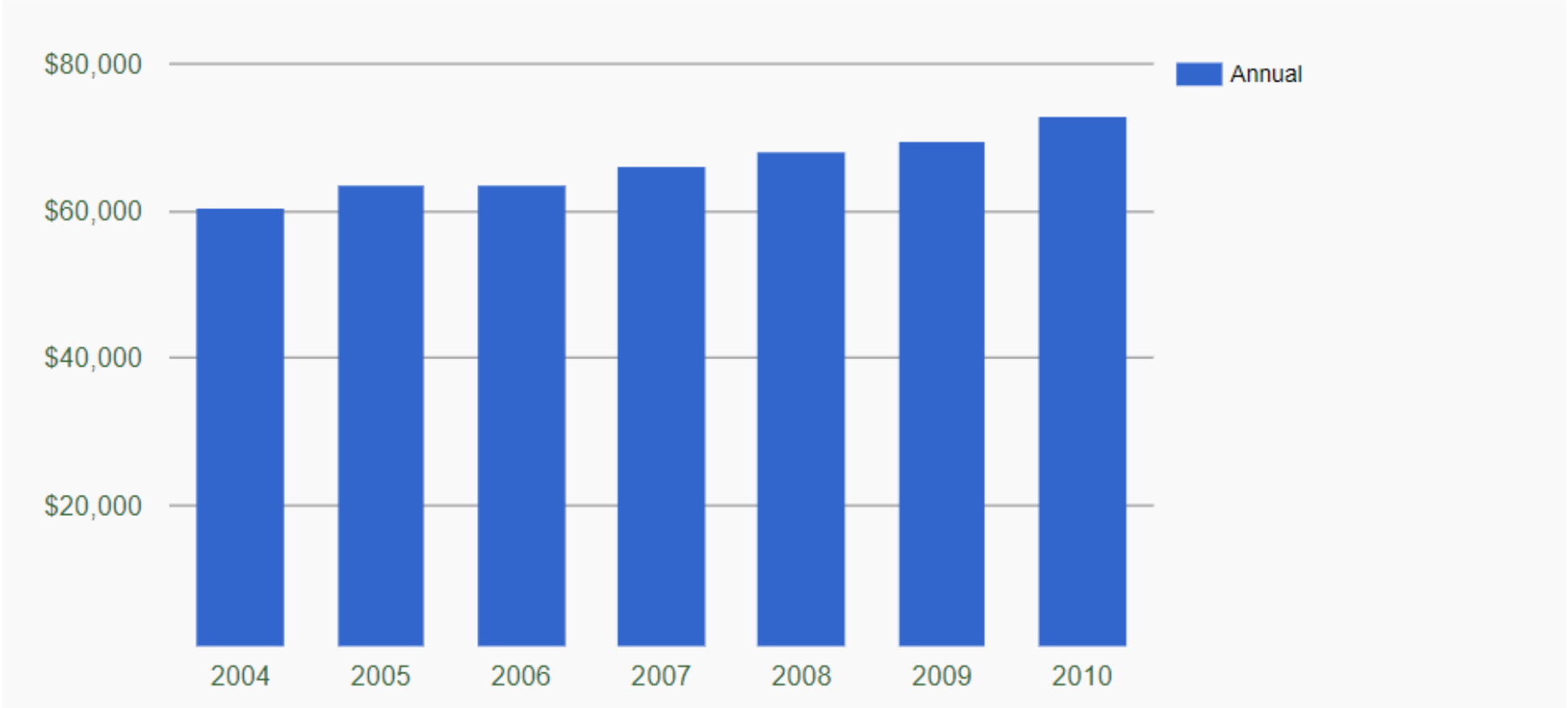
HOURLY RATE

https://www.recruiter.com/salaries/bioinformatics-scientists-salary/

Trends

SALARY TREND

The annual compensation for this career has gone up since 2004. Salaries have increased by an average of 20.87 percent nationwide in that time.



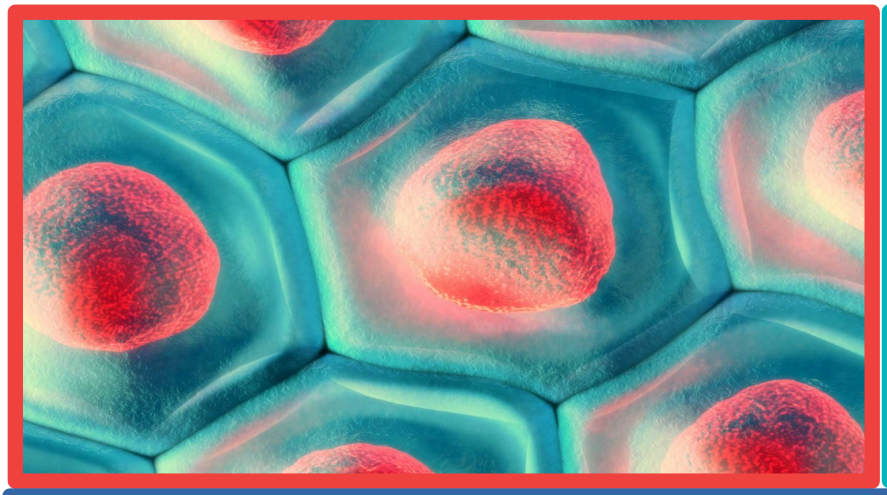
https://www.recruiter.com/salaries/bioinformatics-scientists-salary/

Some Background in Biology

What's a MAJOR Commonality Here?

Organisms Have Genetic Systems!

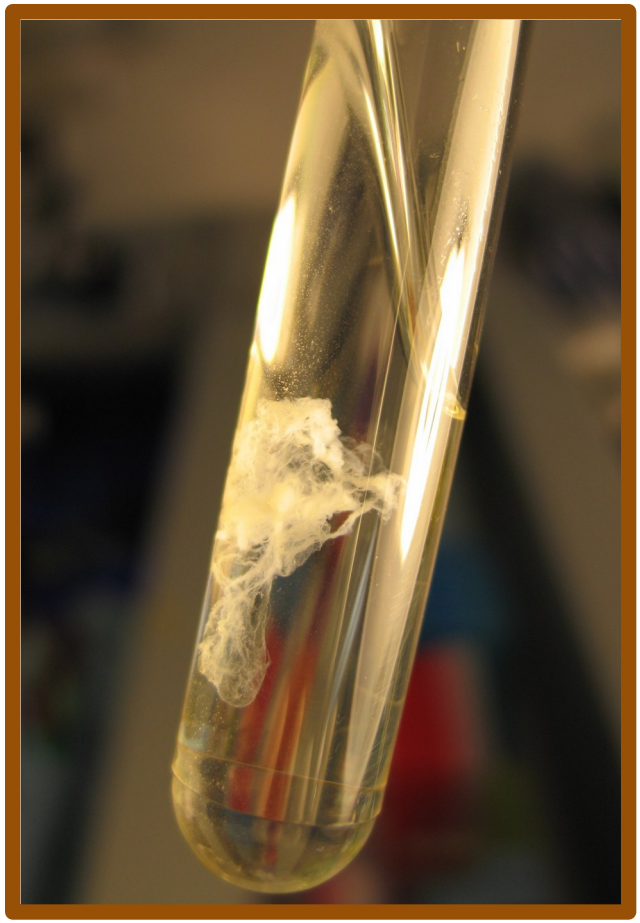
… *And DNA is Often the Genetic Language*



What Does Natural DNA look Like?

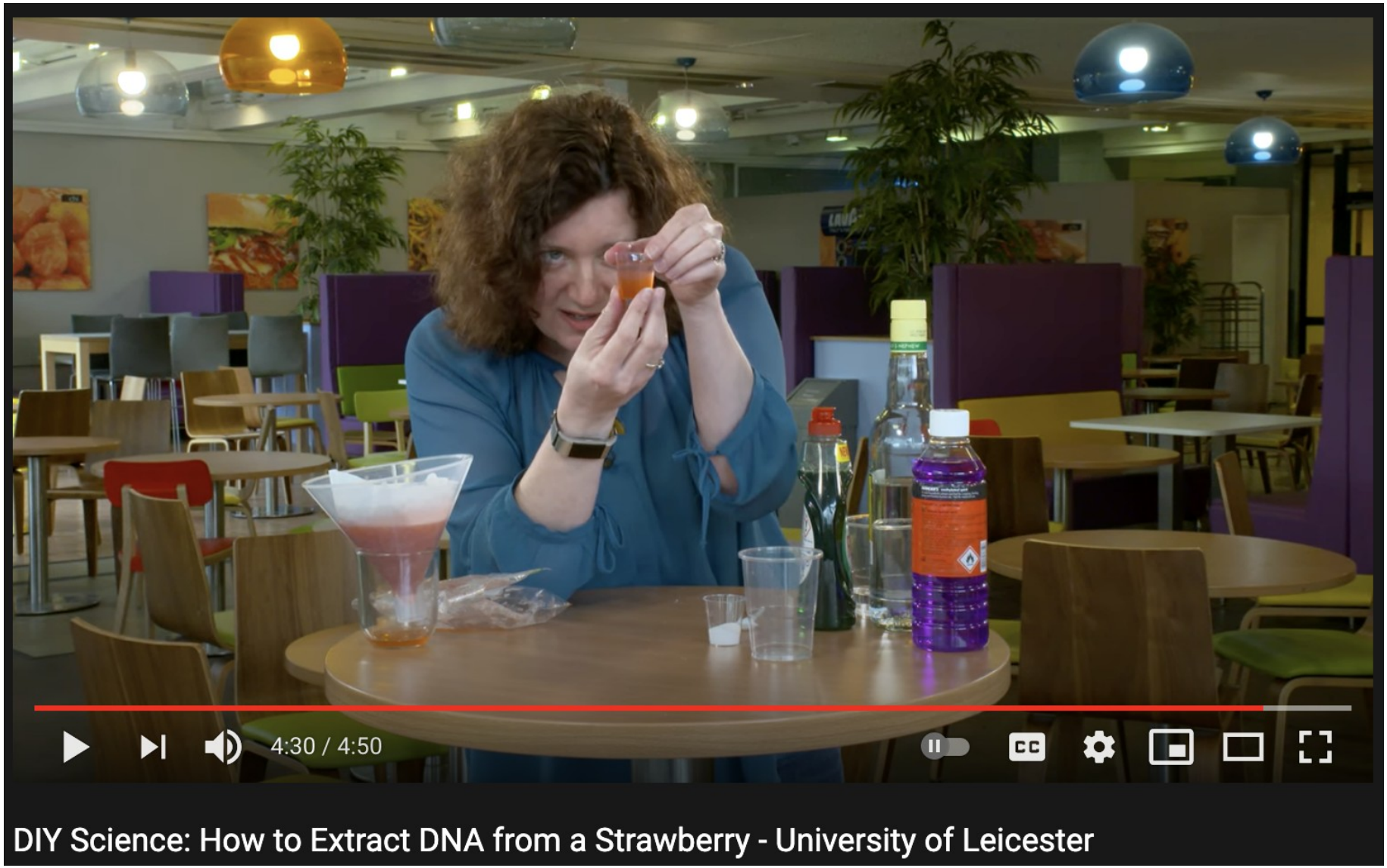
NOT like this!

Like this!



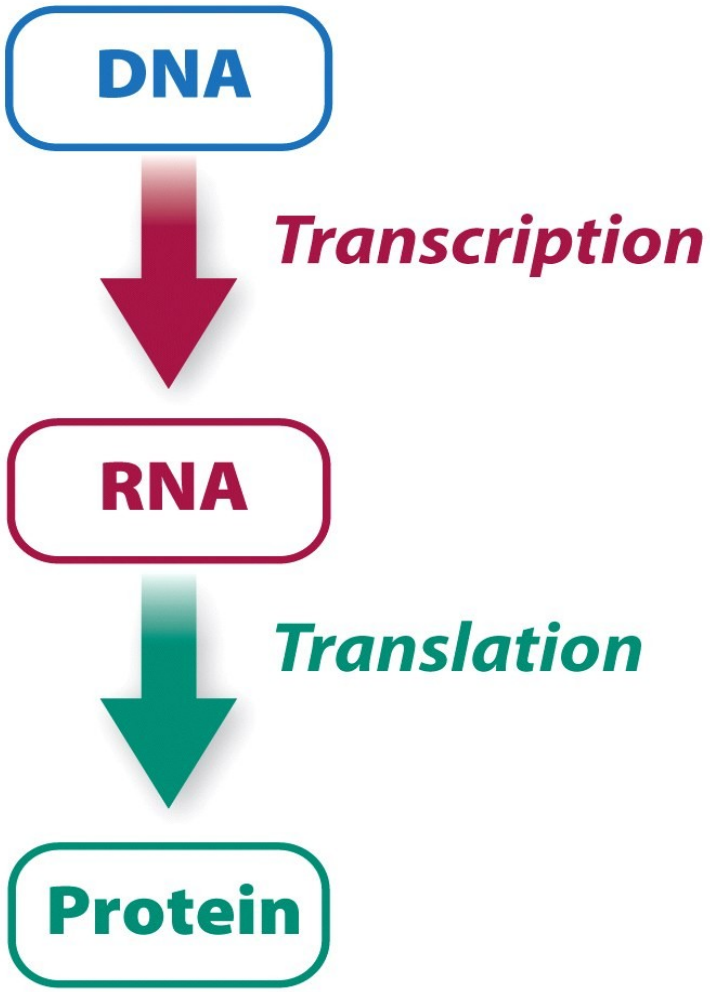
Like this!

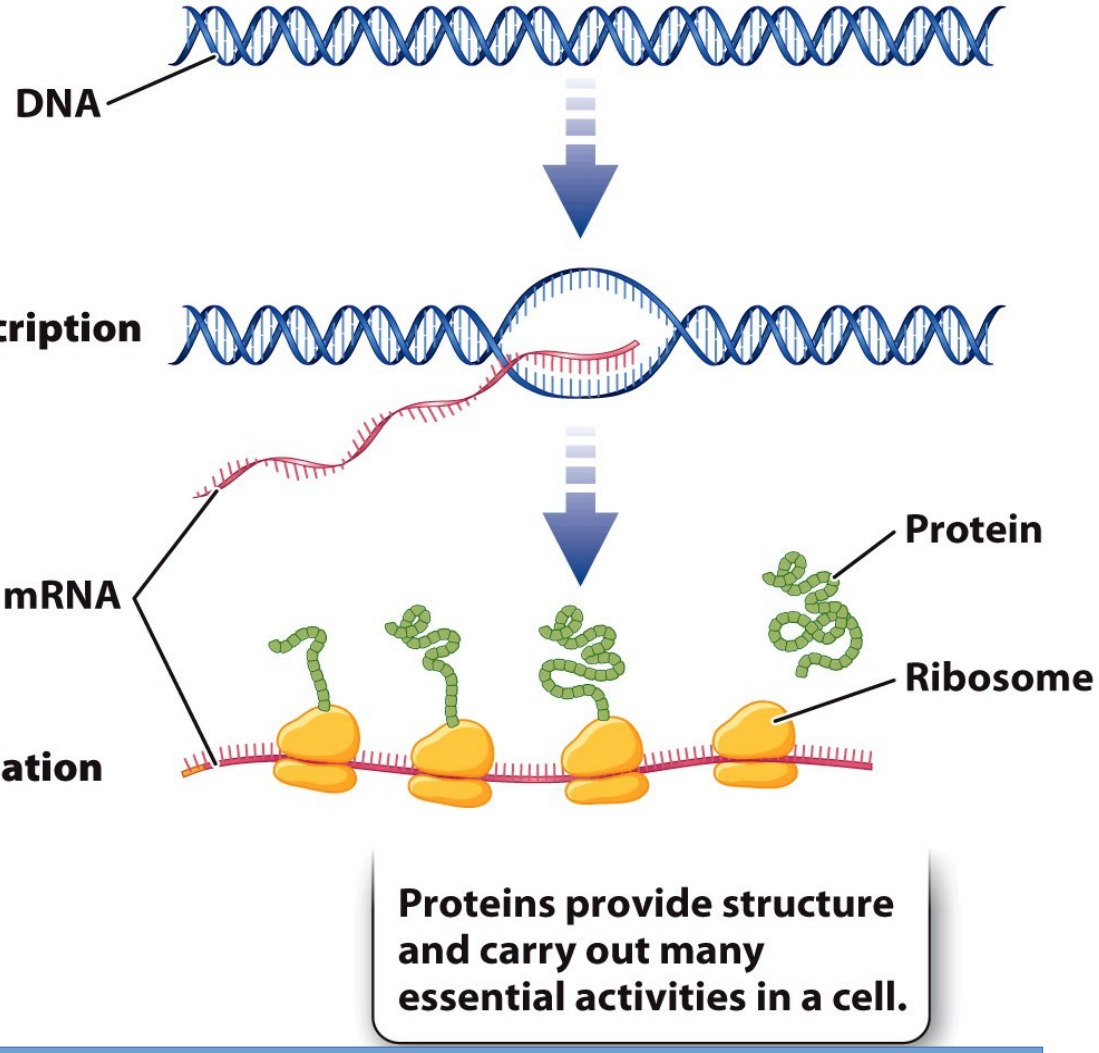
DIY: Strawberry DNA Extraction



Link: https://www.youtube.com/watch?v=JofXXyFZn38

Central Dogma of Molecular Biology

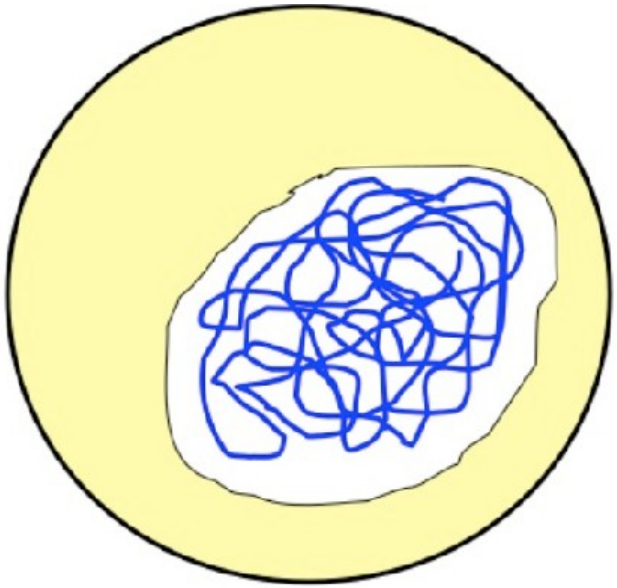


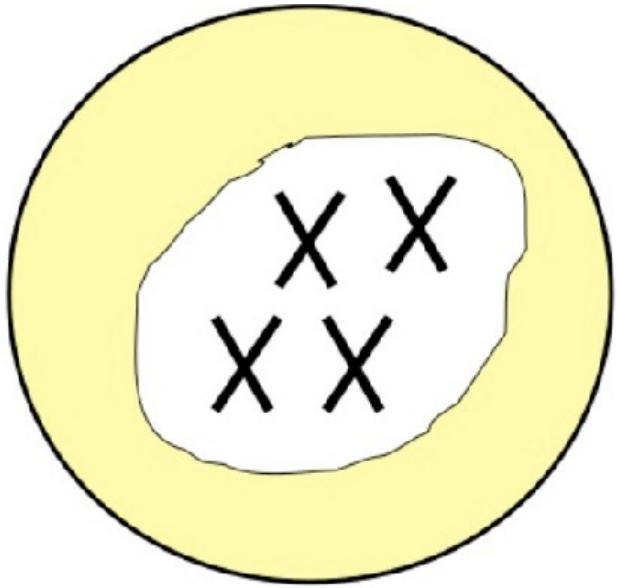


Dogma: *a principle, or set of principles, laid down by an authority as incontrovertibly true.*

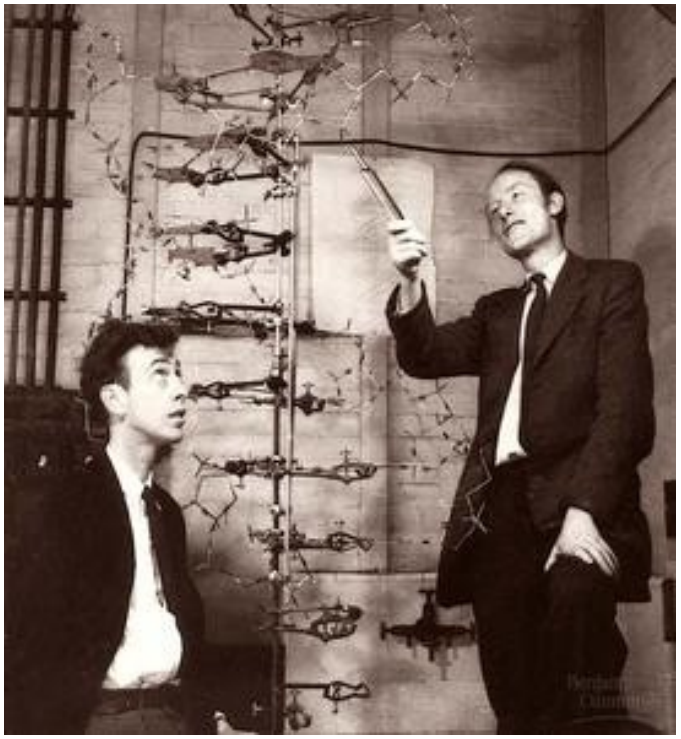
What is DNA?

- Found in the nucleus of a cell in two different structures: chromatin and chromosomes
- Genetic Material (Life's *blueprints*)
- *Written* inherited characteristics (genes)

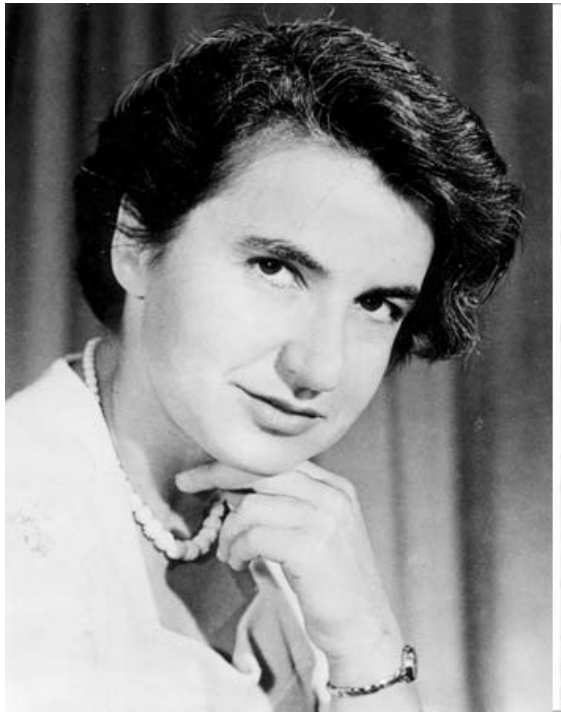


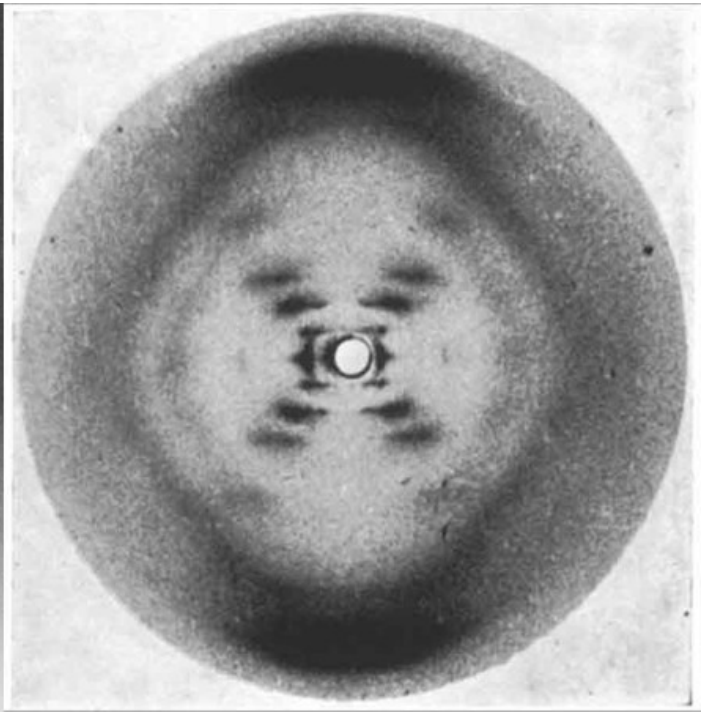


DNA Double Helix:
Discovery of Structure

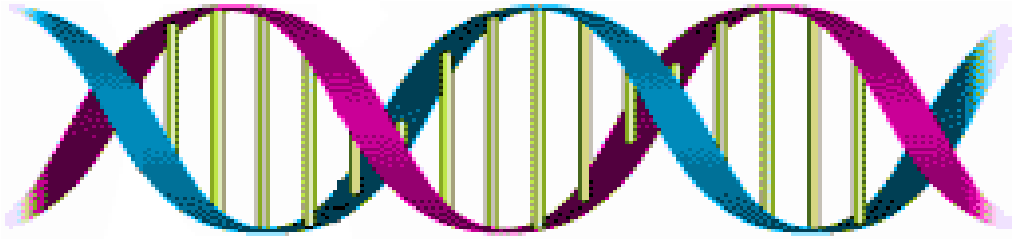


Watson and Crick, 1953

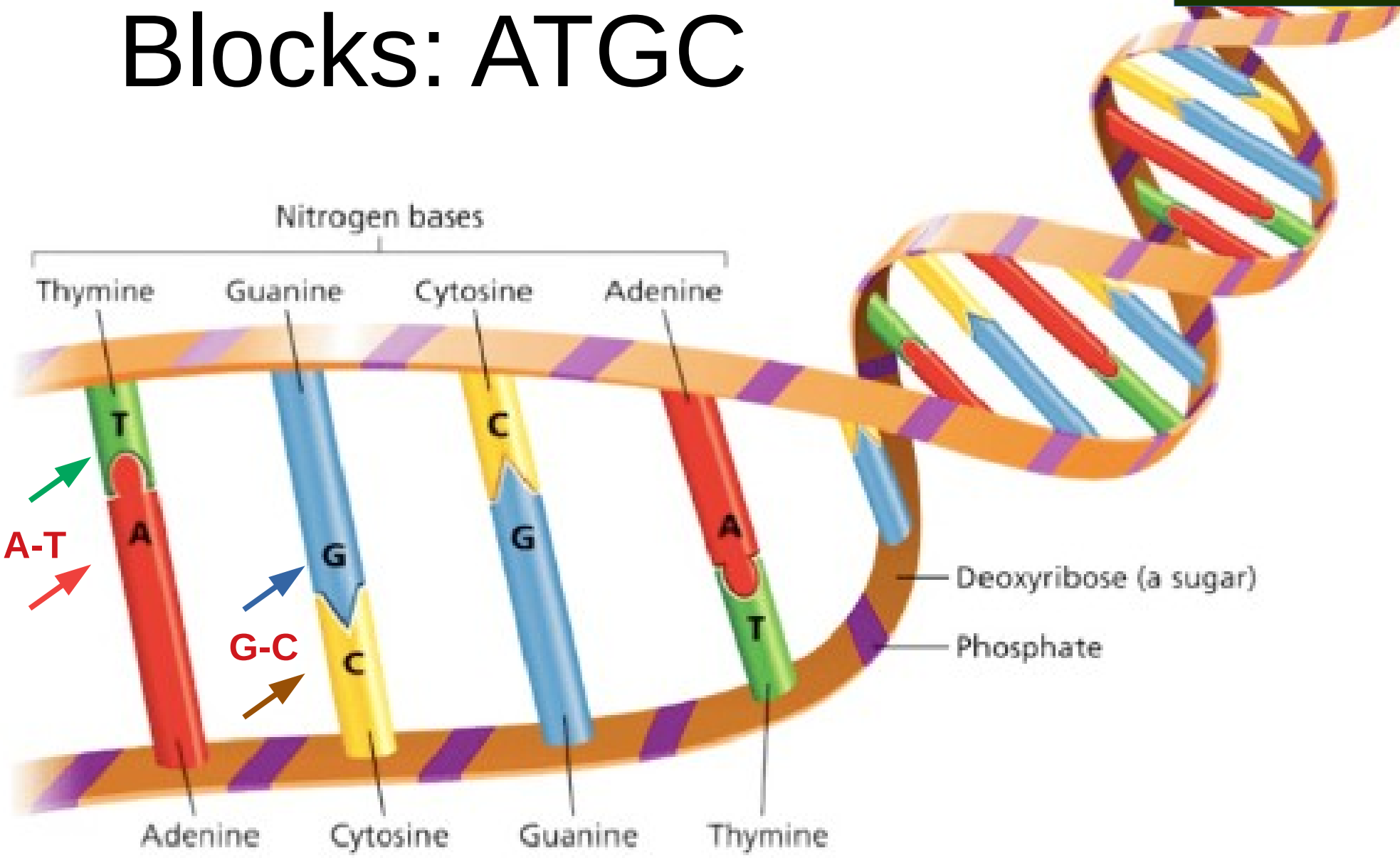




Rosalind Franklin and her data from x-ray crystallography

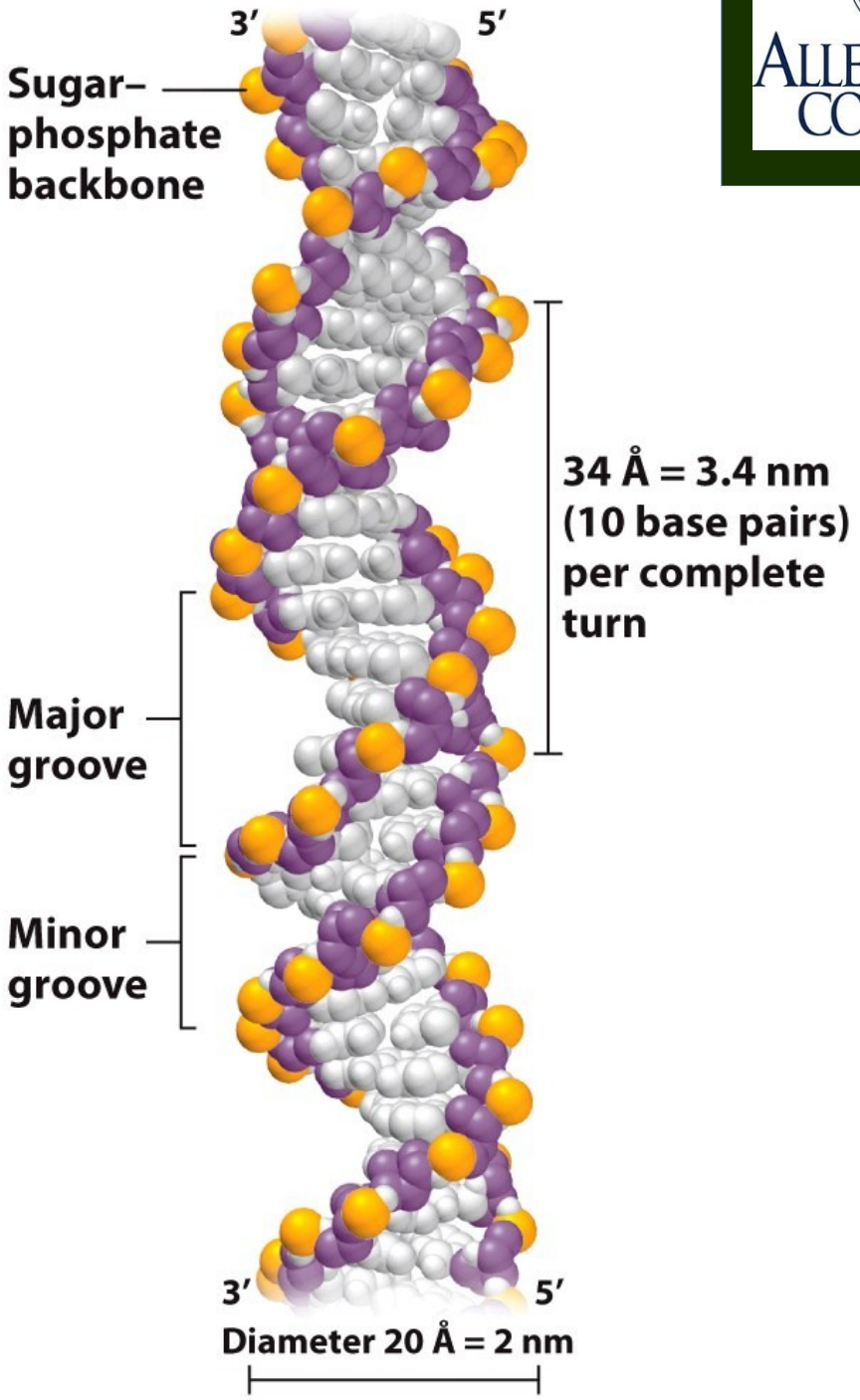


Molecular Building Blocks: ATGC



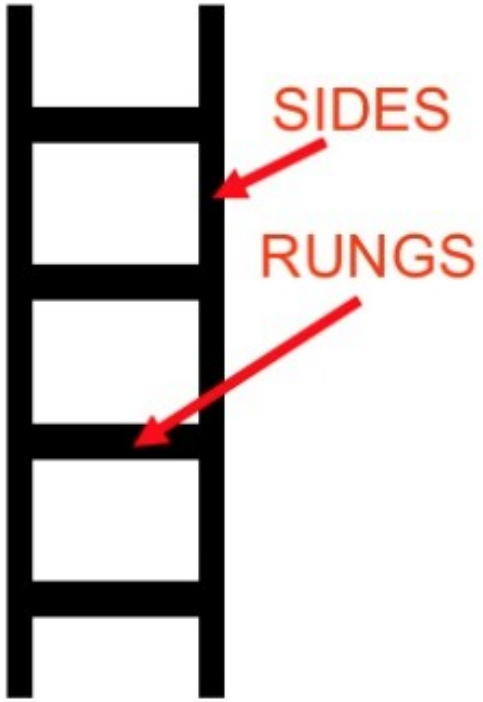
DNA Molecule

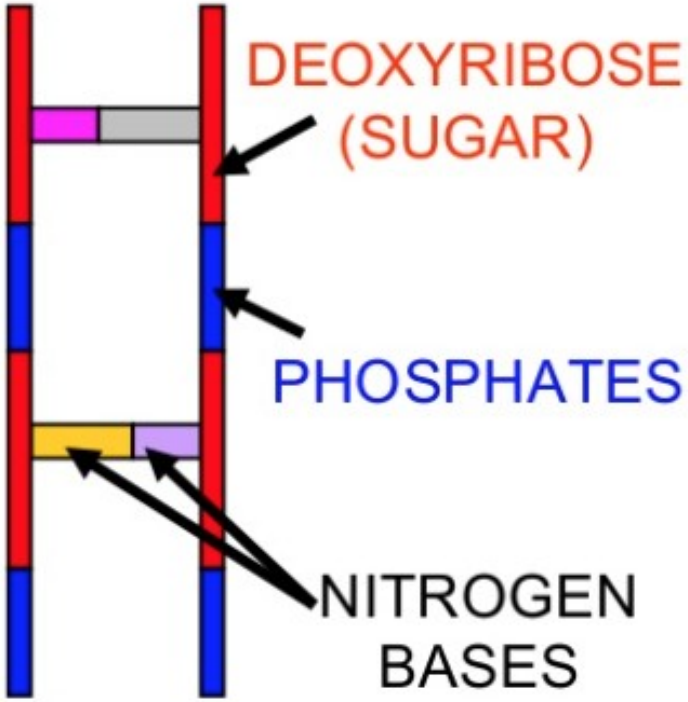
- Double-stranded
- Diameter – 2nm
- Helix
  - Complete turn = 10bp, 3.4nm
  - Major groove
  - Minor groove



DNA Structure

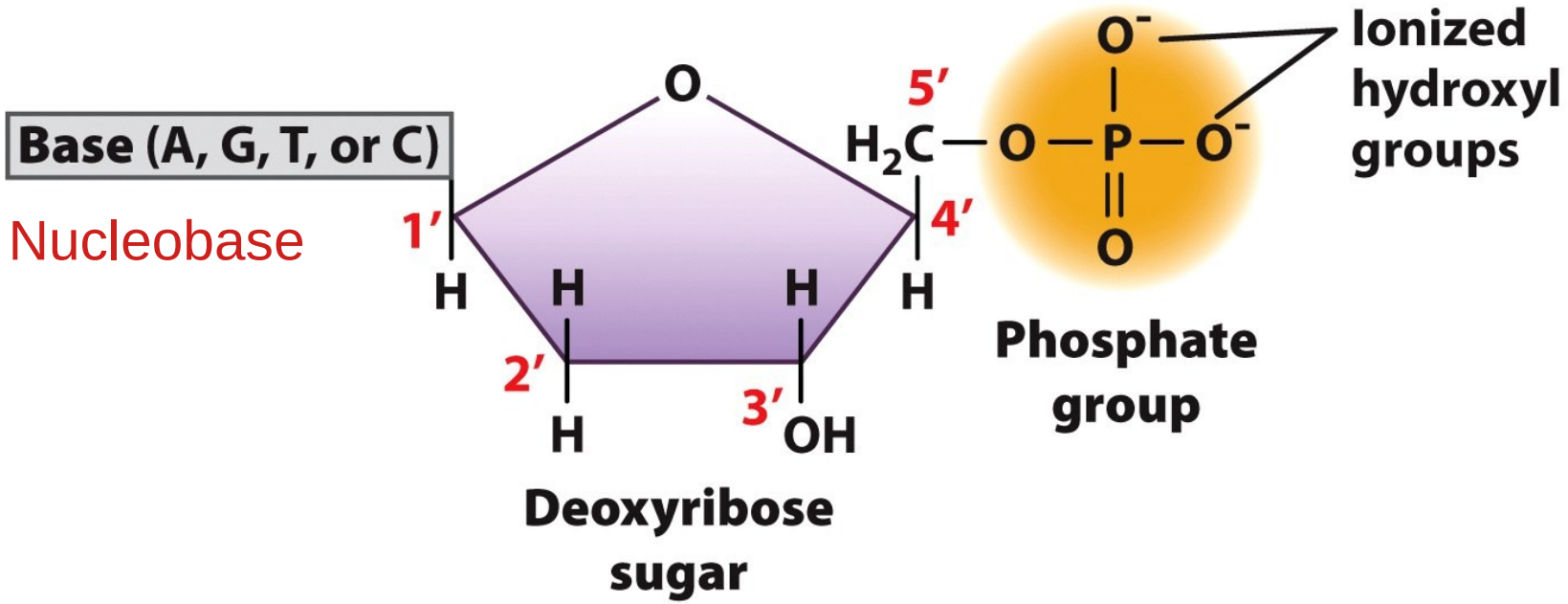
- Formed like a twisted ladder
- There are two sides of the ladder
- Sugar (deoxyribose)
- Phosphates
- Alternating
- Rungs of the ladder
- Nitrogenous bases





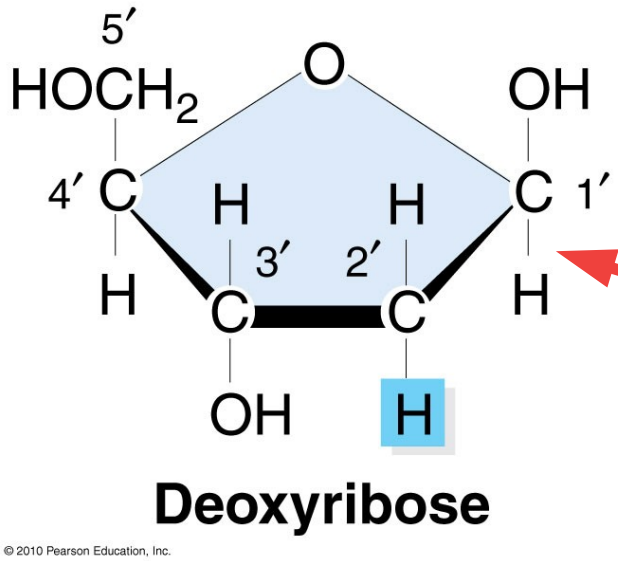
Nucleotides

- The supporting of rungs in the “ladder” (base pairs).
- *Nucleotides: composed of three subunit molecules:*
  - *a nucleobase,*
  - *a five-carbon sugar (ribose or deoxyribose),*
  - *a phosphate group consisting of one to three phosphates.*
  - *The four nucleobases in DNA: guanine, adenine, cytosine and thymine; Note: in RNA, uracil is used in place of thymine.*



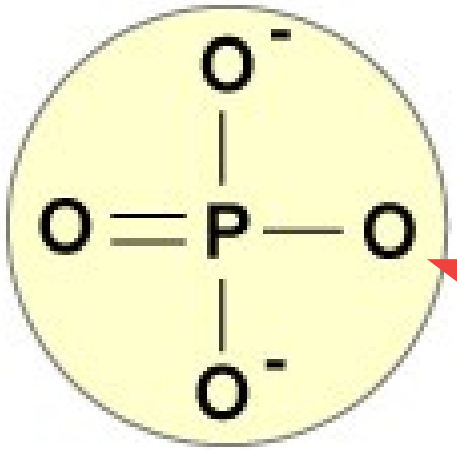
DNA Structure

Deoxyribose **S**ugar



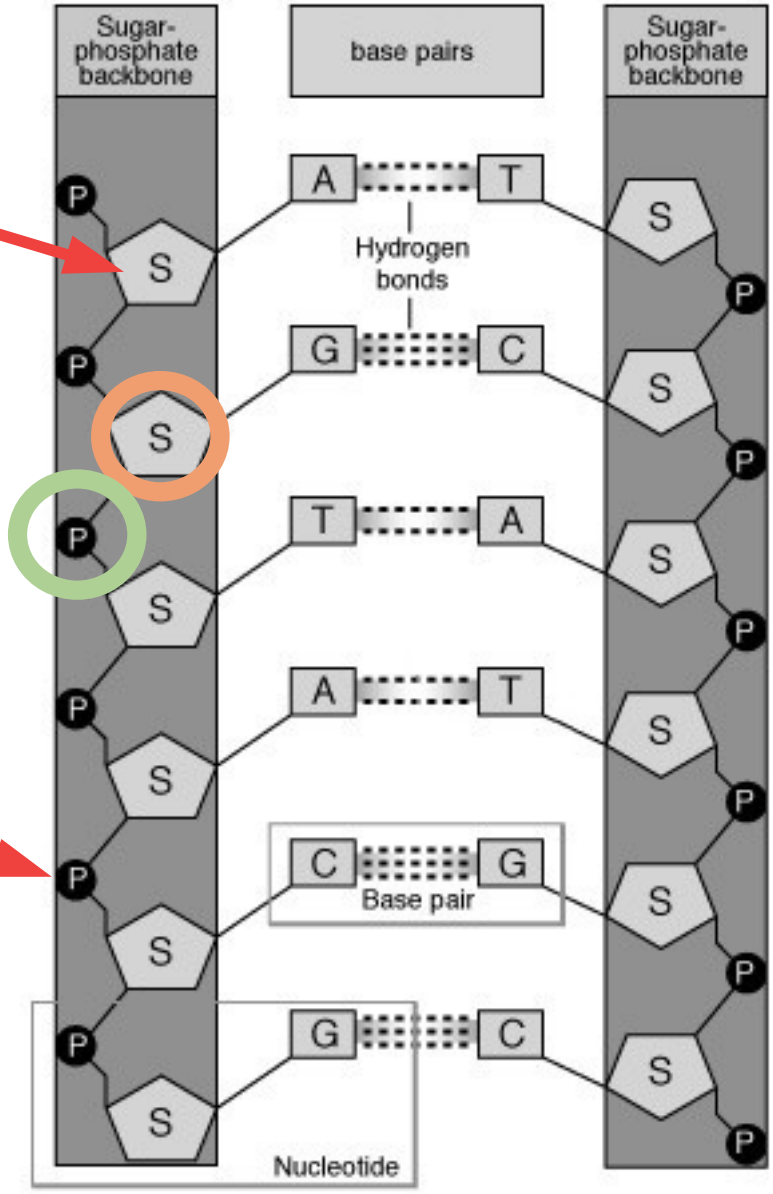
Deoxyribose

Phosphate



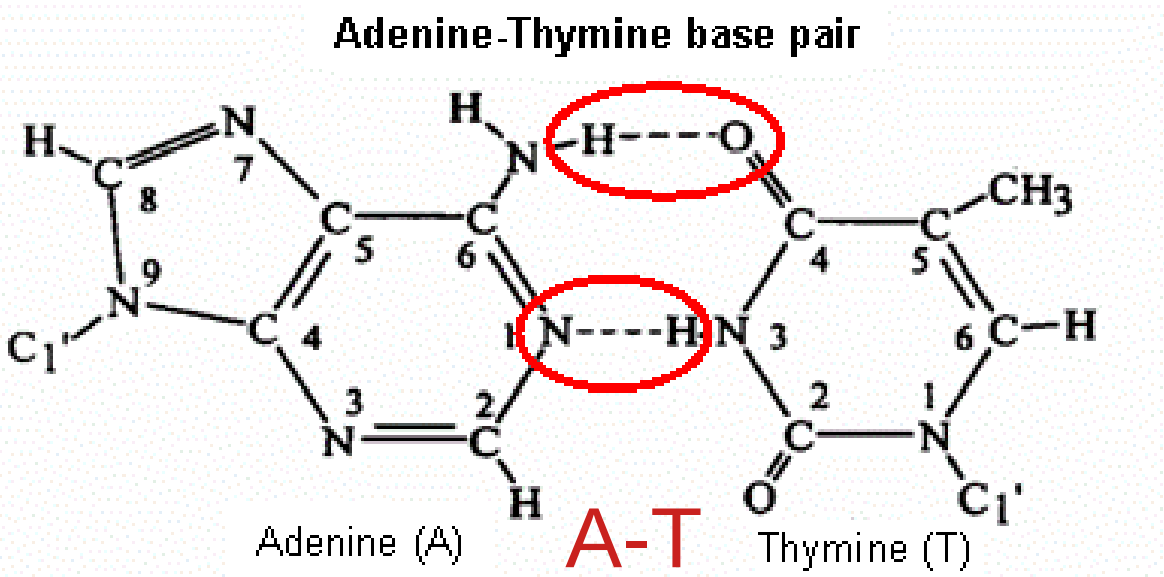
Phosphate group

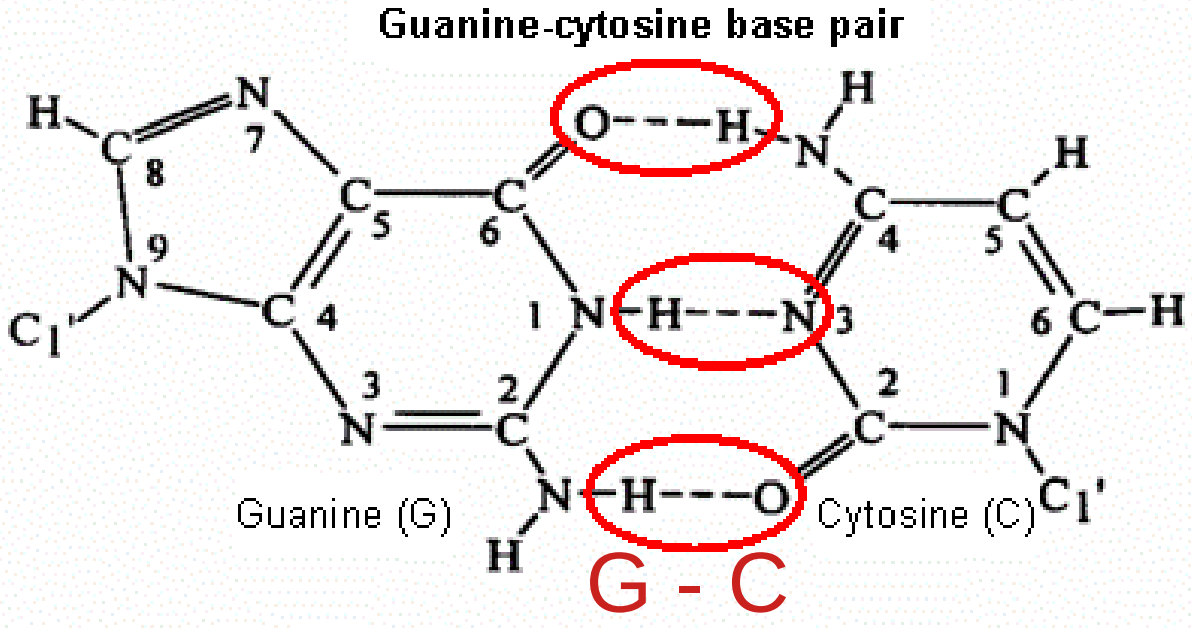
Phosphorus surrounded by oxygens



Base to Base Bonds: How do nitrogenous bases pair?

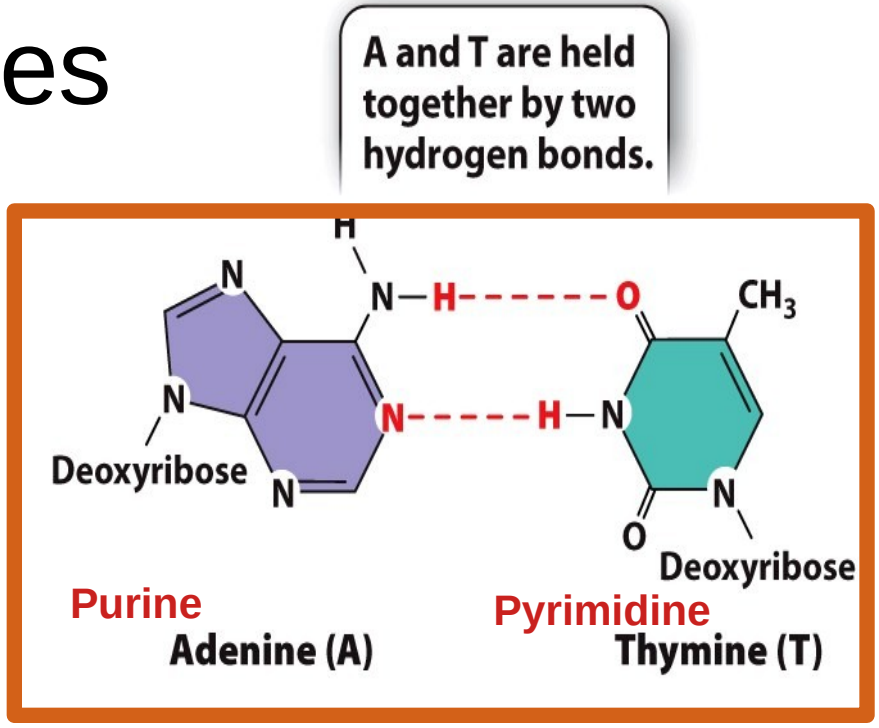
- Base-Specific bonding
- Preserves distance between (DNA's) backbones
- Hydrogen bonds
- Key to replication
- A-T's have **two** bonds
- G-C's have **three** bonds

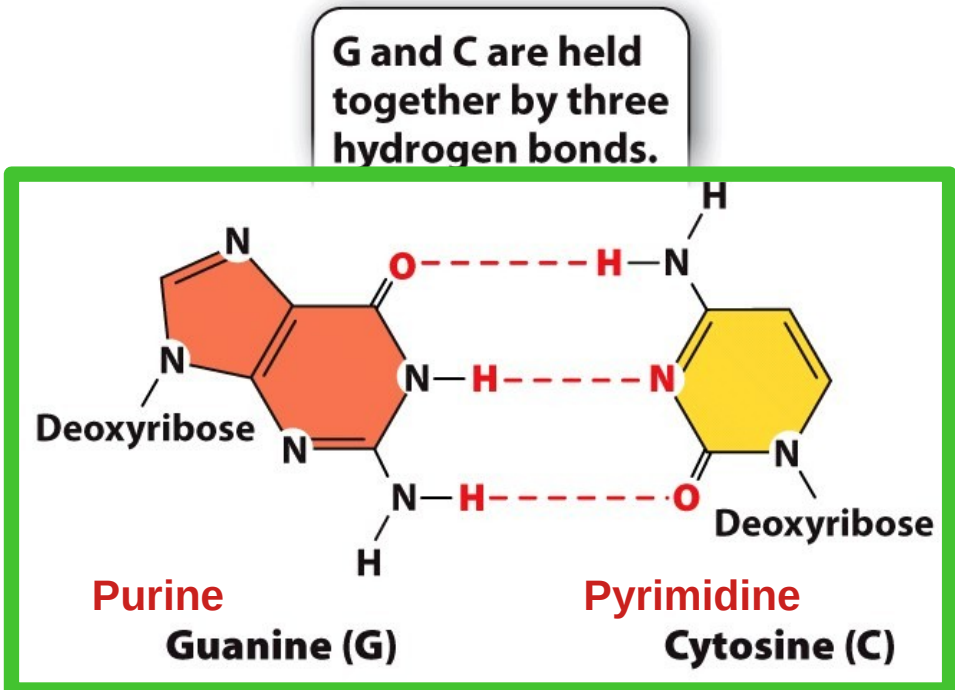




Purines and Pyrimidines

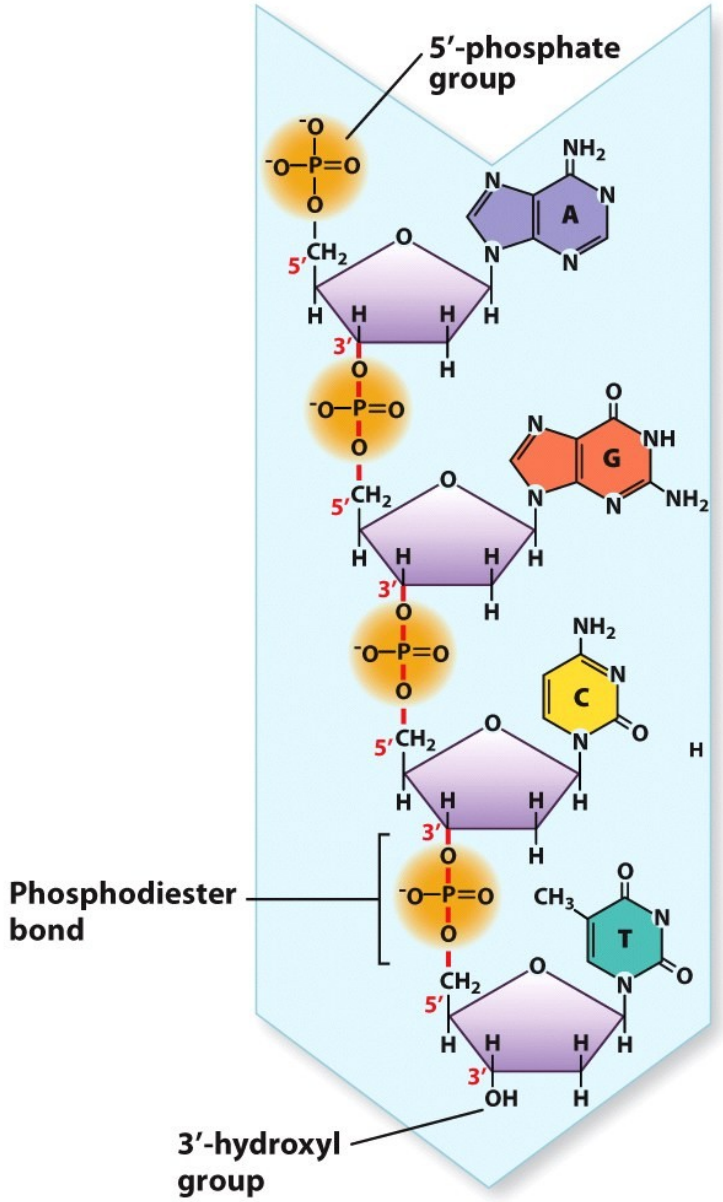
- Purines and Pyrimidines are nitrogenous bases that comprising the two different types of nucleotide bases in DNA and RNA.
- The two-carbon nitrogen ring bases (adenine and guanine) are purines, while the one-carbon nitrogen ring bases (thymine and cytosine) are pyrimidines.
- **Purines**: adenine and guanine
- **Pyrimidine**: thymine, cytosine, and uracil
- Purines include a number of biologically important compounds, such as adenosine, caffeine, uric acid, and the two bases adenine and guanine, which are components of DNA and RNA.





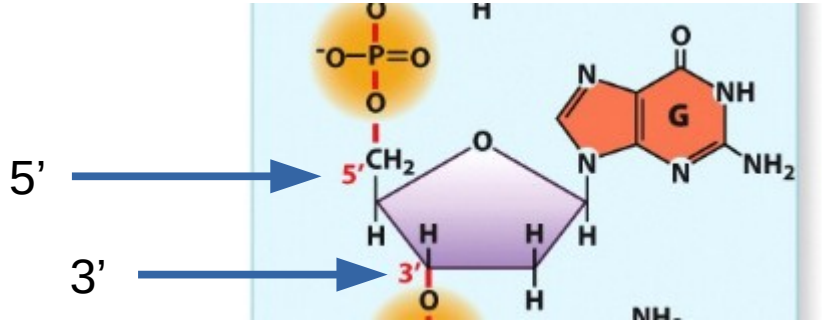
DNA Read in 5’ to 3’ Direction

- Nucleotides are joined by phosphodiester bonds
  - phosphate to sugar
  - covalent bonds
- Polarity
  - 5’ end – phosphate group
  - 3’ end – hydroxyl group
- AntiParallel: DNA read in 5’ to 3’ direction

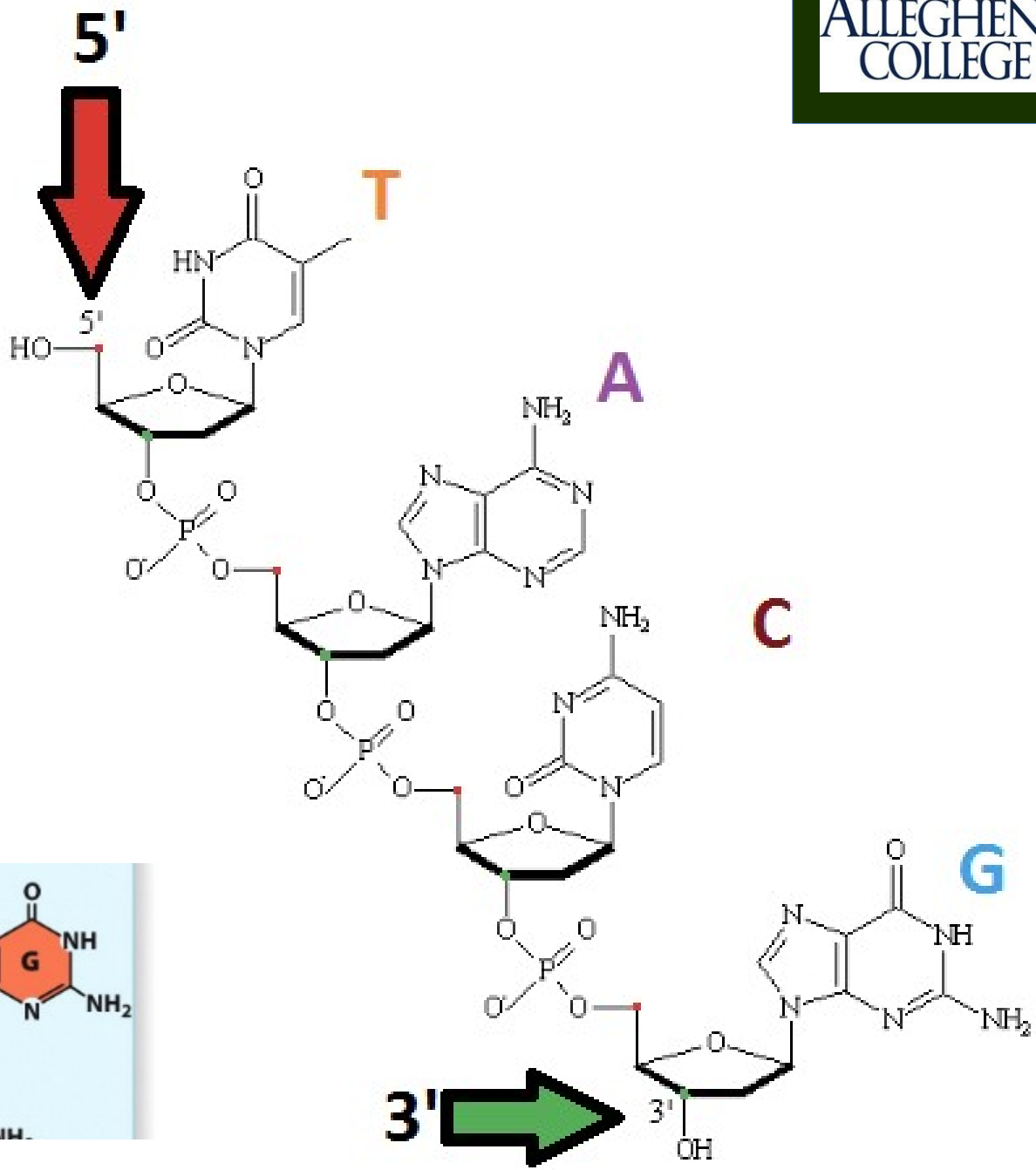


5’ and 3’ Ends?

- A key feature of all nucleic acids is that they have two distinctive ends: **The 5' (5-prime) and 3' (3-prime) ends.**
- This terminology refers to the 5’ and 3’ carbons on the sugar.
- For both DNA and RNA, the 5' end bears a phosphate, and the 3' end a hydroxyl group.

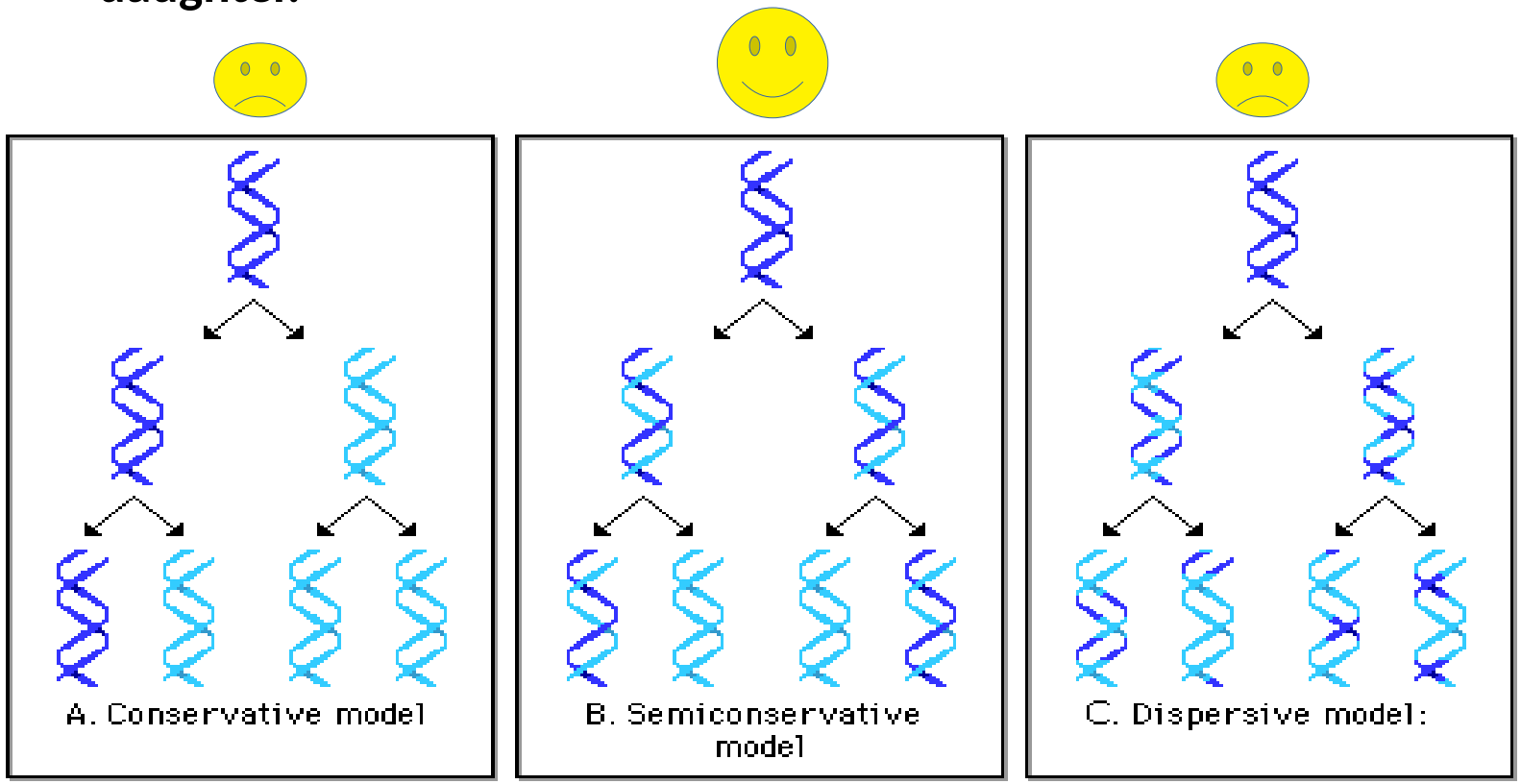


Note the ordering of carbons, hence 5’ to 3’

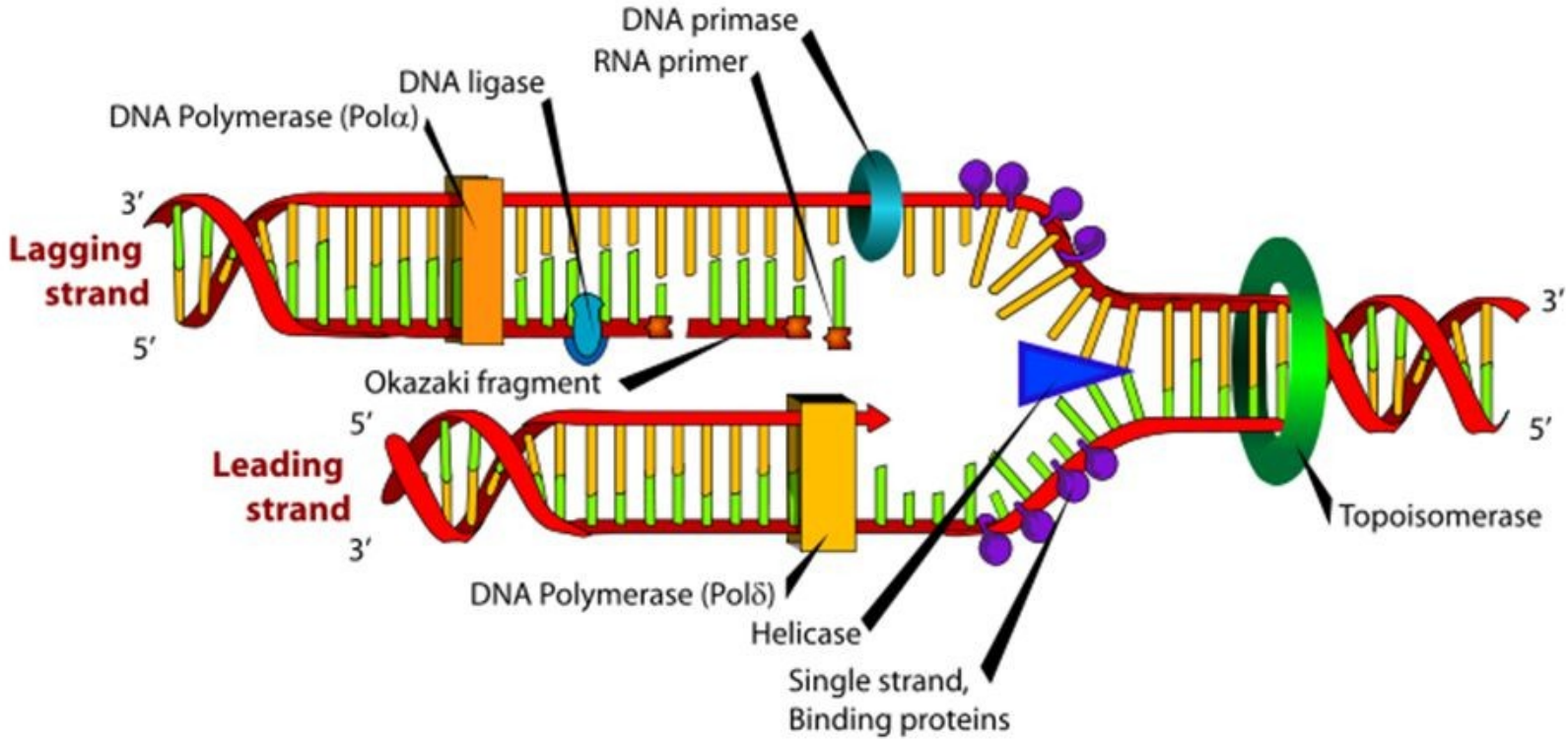


Proposed Mechanisms of Replication Process

1. The two sides of the parent molecule unwind/unzip
2. Daughter strands are synthesized using parent strands as templates
3. Parent/daughter duplex winds back together
   - **Semi-conservative: a 2nd gen helix composed of parental strand and one daughter.**

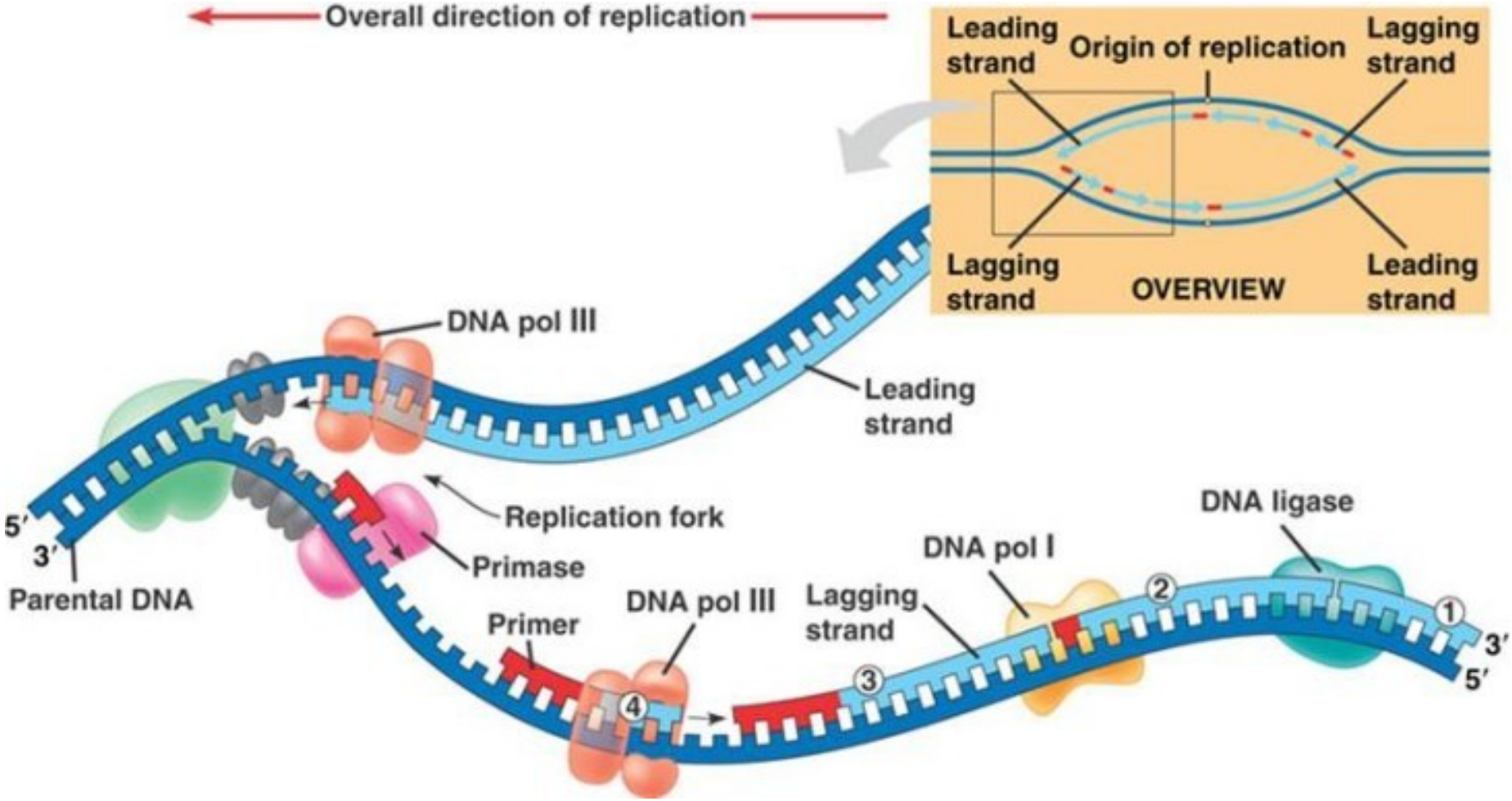


DNA Replication Process



http://www.hhmi.org/biointeractive/dna-replication-schematic

DNA Replication Enzymology



http://www.hhmi.org/biointeractive/dna-replication-basic-detail

http://highered.mheducation.com/sites/0073525324/student_view0/chapter20/dna_replication_fork.html

1. Test Your Understanding

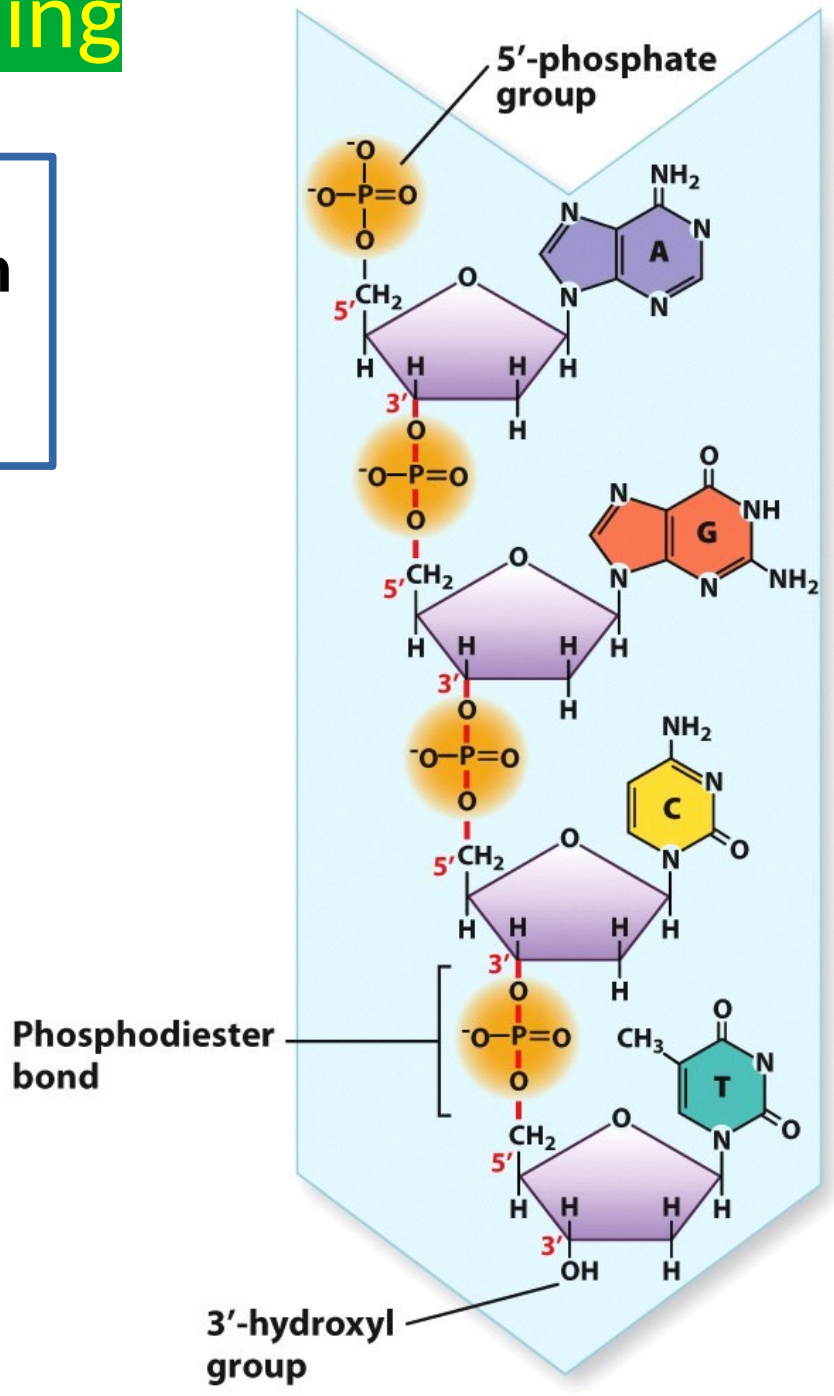
In the DNA sequence 5'–AGCT–3', the phosphodiester linkage between the adenine and the guanine connects:

A. The 2' end of the adenine to the 4' end of the guanine.

B. The 5' end of the adenine to the 3' end of the guanine.

C. The 5' end of the guanine to the 1' end of the adenine.

D. The 3' end of the adenine to the 5' end of the guanine.



1. Test Your Understanding

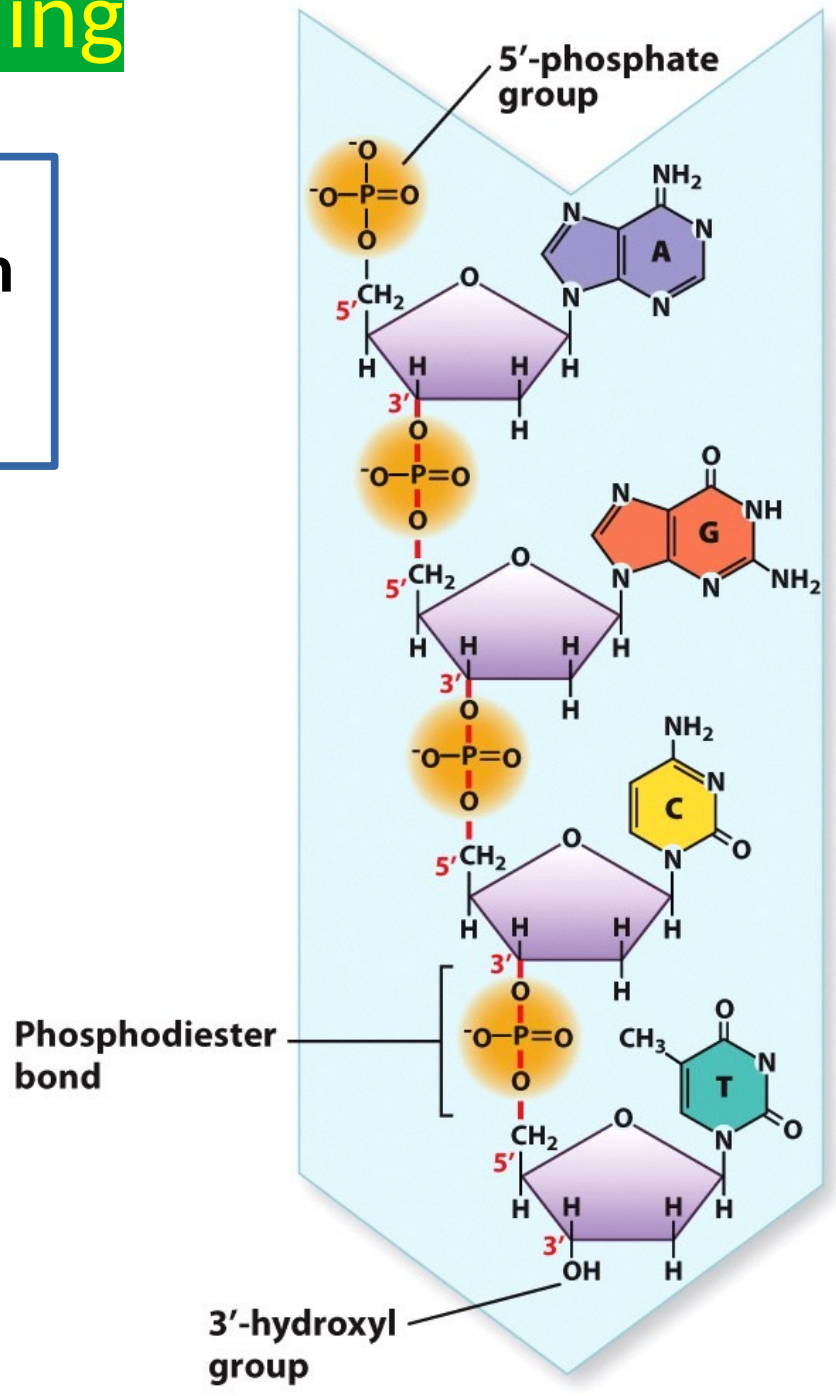
In the DNA sequence 5'–AGCT–3', the phosphodiester linkage between the adenine and the guanine connects:

A. The 2' end of the adenine to the 4' end of the guanine.

B. The 5' end of the adenine to the 3' end of the guanine.

C. The 5' end of the guanine to the 1' end of the adenine.

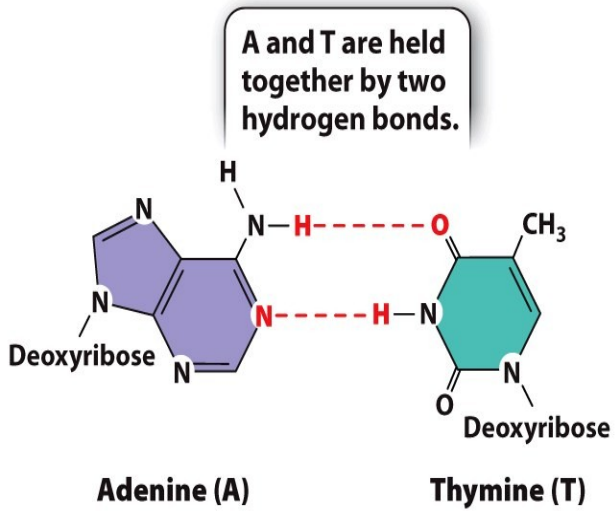
D. The 3' end of the adenine to the 5' end of the guanine.



2. Test Your Understanding

In the DNA of certain bacterial cells, 16% of the nucleotides are adenine. What are the percentages of the other nucleotides in the bacterial helix of DNA?

A. 34% thymidine, 34% guanine, 16% cytosine
B. 34% uracil, 16% guanine, 16% cytosine
C. 16% thymidine, 34% guanine, 34% cytosine
D. 34% thymidine, 16% guanine, 34% cytosine



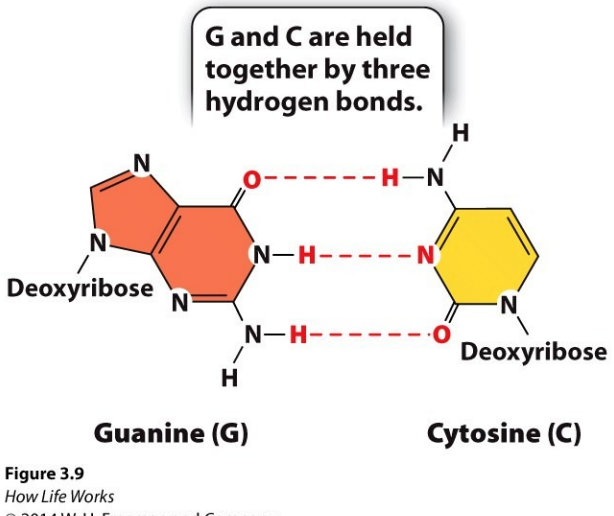
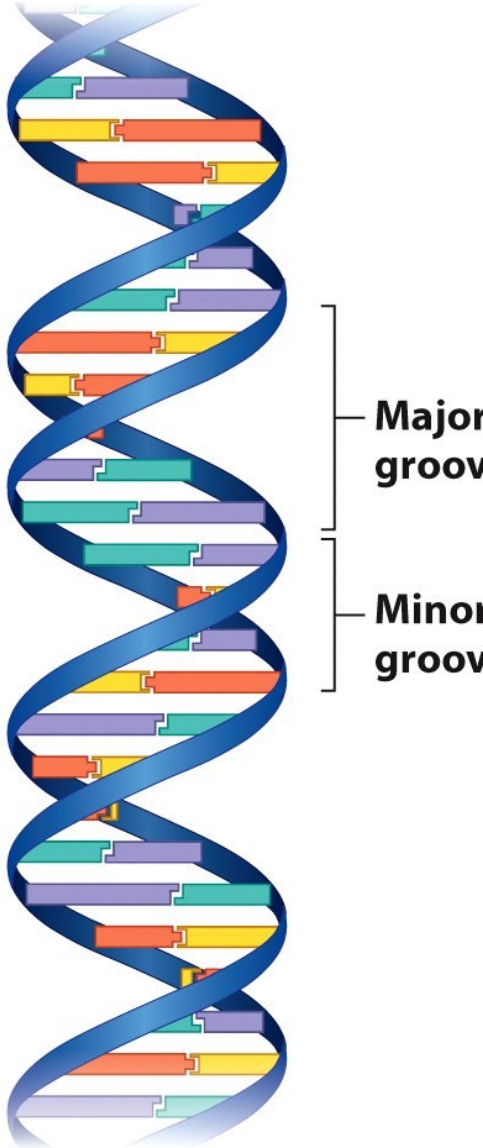


Figure 3.9
How Life Works



2. Test Your Understanding

In the DNA of certain bacterial cells, 16% of the nucleotides are adenine. What are the percentages of the other nucleotides in the bacterial helix of DNA?

A. 34% thymidine, 34% guanine, 16% cytosine

B. 34% uracil, 16% guanine, 16% cytosine

C. 16% thymidine, 34% guanine, 34% cytosine

D. 34% thymidine, 16% guanine, 34% cytosine

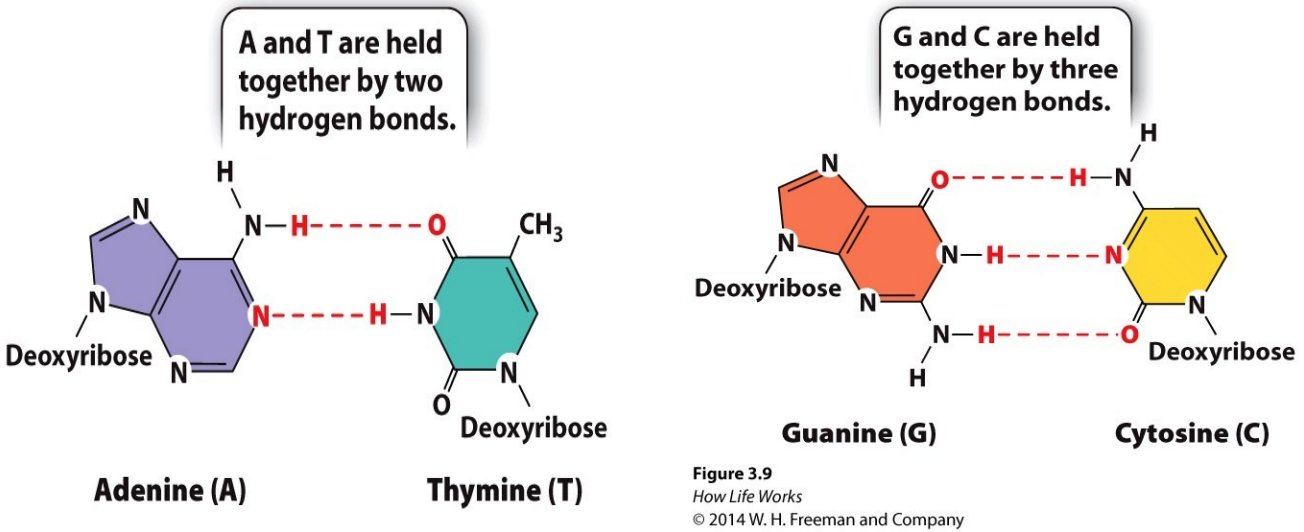
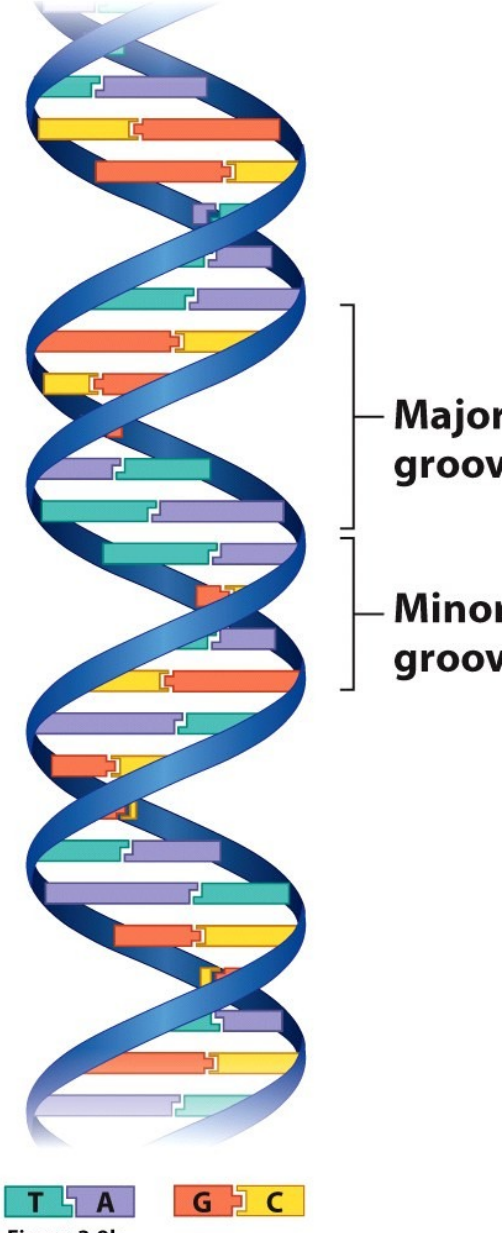


Figure 3.9
How Life Works



3. Test Your Understanding

DNA replicates in a semi-conservative manner. This means:

a) one daughter strand is synthesized as a large fragment while the other is synthesized in smaller fragments, both in the 5'-3' direction
b) every newly formed double-stranded DNA molecule consists of one parental strand and one daughter strand
c) every newly formed double-stranded DNA molecule is comprised of two new daughter strands
d) one daughter strand is synthesized as a large fragment in the 5'-3' direction while the other is synthesized in smaller fragments in the 3'-5' direction

3. Test Your Understanding

DNA replicates in a semi-conservative manner. This means:

a) one daughter strand is synthesized as a large fragment while the other is synthesized in smaller fragments, both in the 5'-3' direction

b) every newly formed double-stranded DNA molecule consists of one parental strand and one daughter strand

c) every newly formed double-stranded DNA molecule is comprised of two new daughter strands

d) one daughter strand is synthesized as a large fragment in the 5'-3' direction while the other is synthesized in smaller fragments in the 3'-5' direction

Supporting Videos

- The Double Helix (Documentary about DNA discovery, 17 mins)
- http://media.hhmi.org/biointeractive/films/Double_Helix.html

- The Chemical Structure of DNA (3 mins)
- http://www.hhmi.org/biointeractive/chemical-structure-dna

- The Structure of DNA (6 mins)
- https://www.youtube.com/watch?v=o_-6JXLYS-k

- The def of 5' and 3' strands (1.5 mins)
- https://www.youtube.com/watch?v=qWZYpHSXvJo