

## Constructing DNA Double Helix with Microsoft Excel

**Subject:** Biology, AP Biology

**Grade:** 9<sup>th</sup> –10<sup>th</sup>, 12<sup>th</sup>

**Rational or Purpose:** In this lesson, students will learn the concept of base-pairing in DNA. They will apply this concept to construct a complimentary strand from a given strand. Since it will be time-consuming to construct a complimentary strand if the given strand is long, the students will learn to use “IF” function in MS Excel to construct a double-stranded DNA.

**Materials:**

Computer with Microsoft Excel

**Lesson Duration:** 50 minutes

**TEKS Objectives:**

§112.43. Biology (c) (2a-d) (6a)

AP Biology

II. Heredity and Evolution

B. Molecular genetics: DNA Structure

**Background:** A DNA double-helix is constructed by two single-stranded DNA. The helices are formed by nucleotides that contain the following bases: adenine (A), cytosine (C), guanine (G), and thymine (T). In a single-stranded DNA, there is no specific order for the combination of the nucleotides. The nucleotides can bond with each other through phosphodiester bonds. If two single-stranded DNA helices need to combine and form a double helix, there are a few constraints that we should be aware of.

The structure of DNA double-helix is supported by the hydrogen-bonding between the bases of nucleotides. They do not hydrogen-bond randomly with other bases, there is a specific requirement for the hydrogen-bond. This is also known as *base pairing*. In base pairing, adenine will hydrogen-bond with thymine, forming an AT base pair. Cytosine will hydrogen-bond with guanine, forming a CG base pair.

If we have a single-stranded DNA sequence, we can easily find the complimentary sequence. If the sequence is not long, it can be done by hand. However, if the strand contains hundreds or thousands of nucleotides, it can take a huge amount of time to complete the complimentary strand by hand. Through the use of Microsoft Excel, the complimentary strand could show within 10 seconds.

**Activity:** This activity will require careful logical thinking when we have to use one general formula to complete the complimentary strand. The activity will begin by introducing the definition of the “IF” function. The first few exercise can be done on paper. As we begin to work on the DNA strands, students will need to use computer to practice and apply the concept.

**Procedure**

1. Follow the steps on the student handout. Students may have difficulties in the very beginning, demonstrating the process can help students.

2. Answer for the bonus question:

=IF(A1="A", "T", IF(A1="T", "A", IF(A1="G", "C", "G")))

**Teacher Notes:**

If A1 is an adenine, the complimentary strand should be a thymine. If A1 is not an adenine, carry out the following IF function.

If A1 is a thymine, the complimentary strand should be an adenine. If A1 is not a thymine, carry out the following IF function.

If A1 is a guanine, the complimentary strand should be a cytosine. If A1 is not a guanine, adenine, and thymine, A1 must be cytosine which makes the complimentary strand to be a guanine.

Name: \_\_\_\_\_

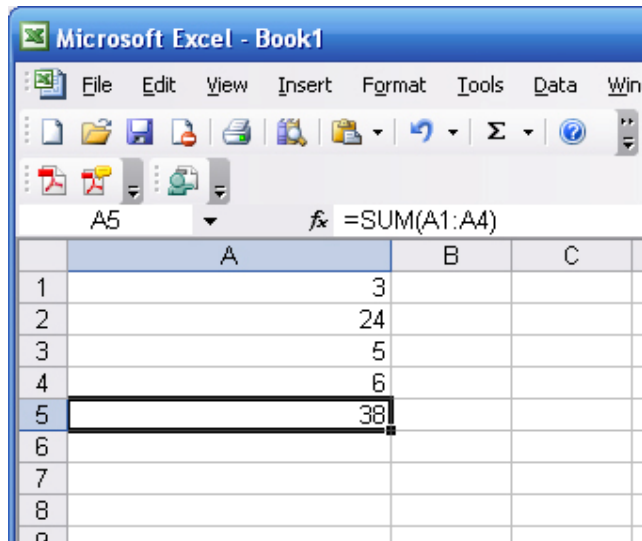
Class Section: \_\_\_\_\_

Date: \_\_\_\_\_

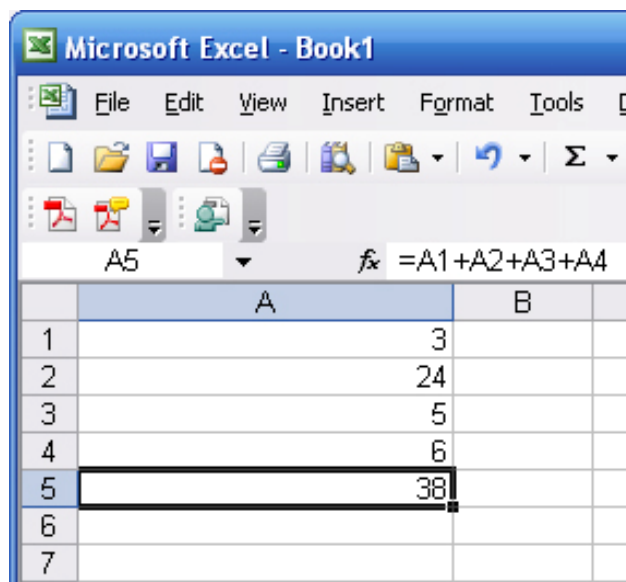
### Constructing DNA double helix with Microsoft Excel

In Microsoft Excel, there are many mathematical and statistical functions. Today you will learn about the usage of the “IF” function.

When you launch MS Excel, you will see a spreadsheet with many cells. Whenever you enter a function, you must start with an equal sign “=”. See the following picture:



Let's analyze the function. In cell A5, you see “38” in the cell. It is calculated by a function “=SUM(A1:A4)”. This function means cell A5 is equal to the sum between A1 and A4, which is  $3 + 24 + 5 + 6 = 38$ . You may ask, will “=A1 + A2 + A3 + A4” give the same result? Look at the next example:

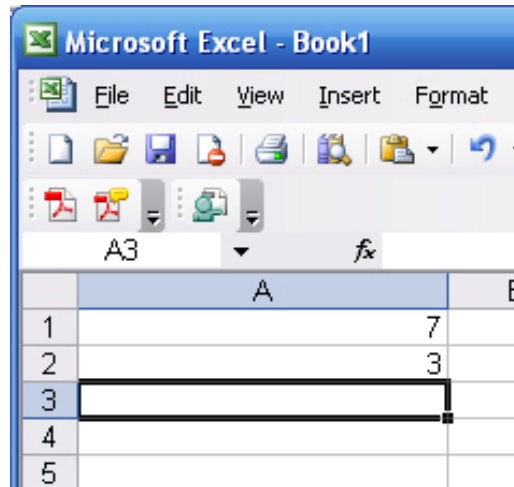


Name: \_\_\_\_\_

Class Section: \_\_\_\_\_

Date: \_\_\_\_\_

It yields the same result. Based on this example, you can use MS Excel to conduct simple calculations. Consider this example:



You have two numbers in A1 and A2 respectively. What will you enter in A3 in order to get the product of A1 and A2? (Multiplication in Excel is represented by the symbol “\*”).

A3 = \_\_\_\_\_

Remember, whenever you enter a function in Excel, the function must start with an equal sign.

Let's start learning about the “IF” function now.

From your background knowledge of base-pairing, you know that adenine can only pair with thymine, and cytosine can only pair with guanine. Given a short single-stranded DNA sequence, you can simply find out the sequence of complimentary strand by hand.

Given DNA Strand	-----	Complimentary Strand
A		T
T		A
T		A
G		C
C		G
G		C
T		A

If you have 100 or 1000 bases in a given DNA strand, it can be very time consuming to find out the complimentary strand by hand. In this case, scientists can rely on computers to do this job for us.

First you need to know the structure of the “IF” function.

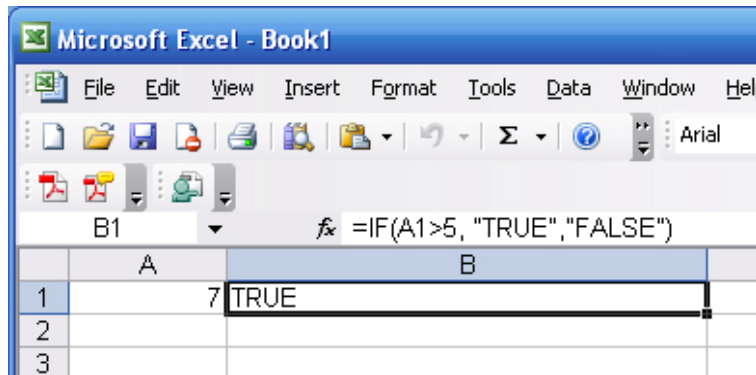
Name: \_\_\_\_\_

Class Section: \_\_\_\_\_

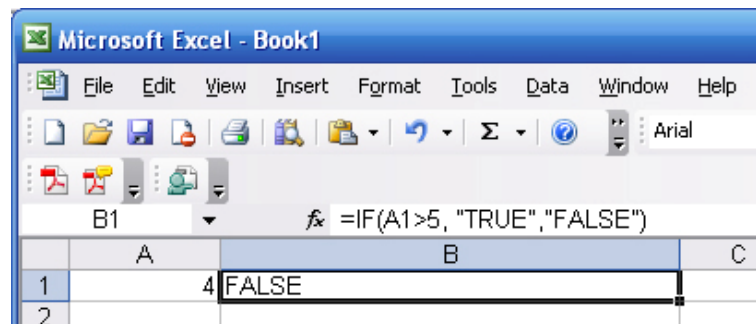
Date: \_\_\_\_\_

= IF (logical\_test,[value\_if\_true],[value\_if\_false])

The “logical\_test” is a mathematical statement. If the statement is true, the value of [value\_if\_true] will be shown. If not, the value of [value\_if\_false] will be shown instead. It may sound a little confusing right now. Take a look at the following example:



In the example, we test if cell A1 is larger than 5. If it is true, then “TRUE” will be displayed in the cell; if it is not true, “FALSE” will be displayed in the cell. You should always take a closer look at the function and make sure you understand the meaning of the function. Remember to put quotation marks between non-numerical values, such as TRUE and FALSE.



In the above example, you can see that the function stays the same but the value of cell A1 is now 4. Since 4 is NOT larger than 5, the statement becomes false. Therefore “FALSE” is shown in cell B1.

### *Apply “IF” Statements to DNA*

Understanding how the “IF” function works, you can apply it to find a complimentary strand of DNA. For example, since there is a thymine in the given strand, you know that the corresponding base at the complimentary strand will be adenine. You can use Excel to write a function to help us.

You could come up with the “logical\_test” and the “[value\_if\_true]”, but what will be the “[value\_if\_false]”? At this stage, it is very difficult to answer this question because we have 4 different bases in total. To make it simpler, let’s consider a strand

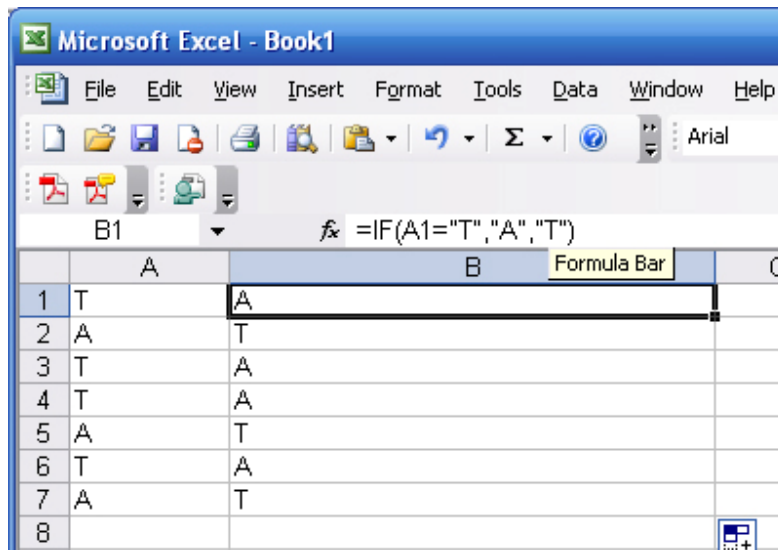
Name: \_\_\_\_\_

Class Section: \_\_\_\_\_

Date: \_\_\_\_\_

consisting only two bases: adenine and thymine. That means if the given has adenine, the complimentary strand must be adenine at the corresponding location. In this case, our function will be easy to write because the base at the complimentary strand must be either adenine or thymine. The function should look like this:

= IF (A1= "T", "A", "T") OR = IF (A1= "A", "T", "A")



One advantage of using Excel is that you can copy and paste the same function to the cells below. They will automatically conduct the test and give results immediately. If the sample strand has only cytosine and guanine, simply change the symbol of the base and follow the same procedure.

*Task:* What if the given single-stranded DNA sequence is: CGCCGGGCCGCGCGGGC? Can you come generate the complimentary strand by using the IF function?

*Bonus:* In reality, a strand of DNA always contains all four bases. How would you write the "IF" function if you are given the following strand: CGTTGCAGCGATA?