Describing the Structure of DNA

Adapted from the University of Buffalo Center for Case Study Teaching in Science & B. Franckowiak

In the autumn of 1952, while the weather in London, England, was cooling down, the race to discover the structure of DNA was heating up inside two rival laboratories: the Cavendish Lab at Cambridge University and the laboratory of J.T. Randall at nearby King's College. DNA had recently been identified as the hereditary material that organisms use to transmit genetic traits from one generation to another. This finding is largely attributed to the work of two groups of American scientists, Oswald Avery, Colin MacLeod, and Macyln McCarty of the Rockefeller Institute, and Alfred Hershey and Martha Chase of Cold Spring Harbor Laboratory, though many important experiments conducted by other scientists were crucial to developing this understanding. The next great challenge was to identify the structure of DNA so that we could understand how the information contained in DNA is copied and passed from cell to cell, and how that information is translated into the instructions for building the proteins necessary for life.

The groundbreaking research of the Cavendish and Randall laboratories describing the structure of DNA was published in a series of papers in the journal Nature on April 25, 1953 [Wilkins et al., 1953; Franklin et al., 1953]. These would become known as some of the most important scientific papers ever published, as well as a symbol of great controversy involving allegations of misogyny and ethics violations among scientists who would go on to win a Nobel Prize.

The Original Research Paper

Watson, J.D. and Crick, F.H. (1953) Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature.* 1953 Apr 25;171(4356):737-8.

This original paper was a mere single page that served, to assert that the pair had "won" the battle to sort out the structure of DNA but not actually convey much information.

One month after the original paper, the full "full details" promised by Watson and Crick were published:

The Second Research Paper

Watson, J.D. and Crick, F.H. (1953) Genetical Implications of the Structure of Deoxyribonucleic Acid. *Nature.* 1953 May 30;171(4361):964-967.

In this assignment, we will examine this 2nd paper. We'll examine the included figures first, then walk through the text of the paper.

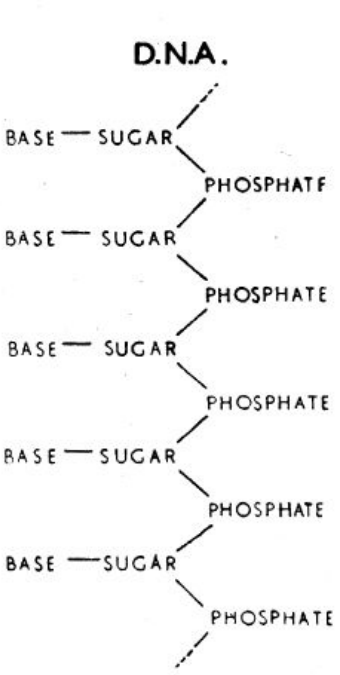

Fig. 1. Chemical formula of a single chain of deoxyribonucleic acid

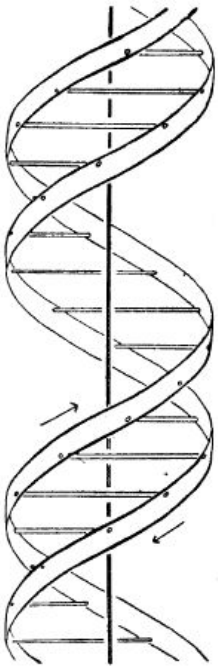

Fig. 2. This figure is purely diagrammatic. The two ribbons symbolize the two phosphate-sugar chains, and the horizontal rods the pairs of bases holding the chains together. The vertical line marks the fibre axis

What does Figure 1 (Fig.1) show the reader about the structure of DNA?

Define NUCLEOTIDE and highlight one nucleotide on the diagram in Figure 1.

What is the purpose of Figure 2 (Fig. 2)?

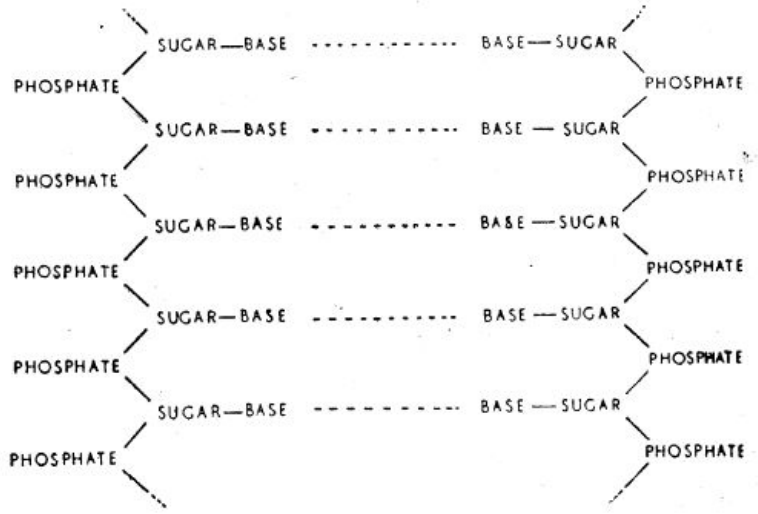

Fig. 3. Chemical formula of a pair of deoxyribonucleic acid chains. The hydrogen bonding is symbolized by dotted lines

Highlight all of the hydrogen bonds in Figure 3 (Fig 3). What does Figure 3 tell the reader about the structure of DNA?

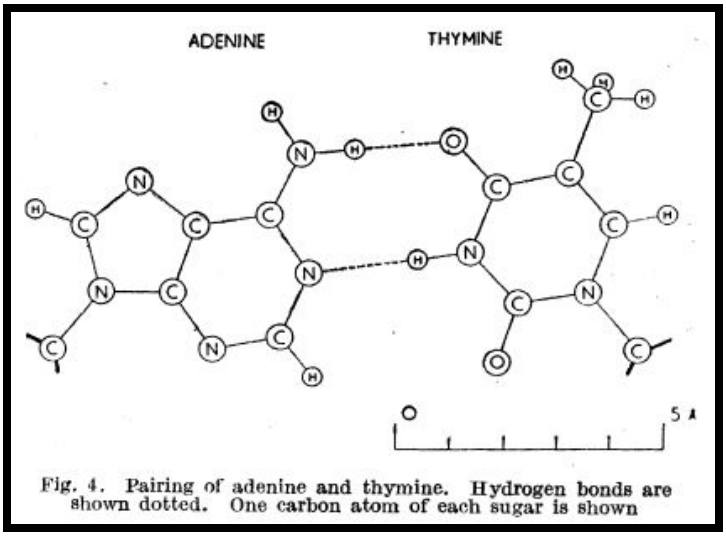

Fig. 4. Pairing of adenine and thymine. Hydrogen bonds are shown dotted. One carbon atom of each sugar is shown

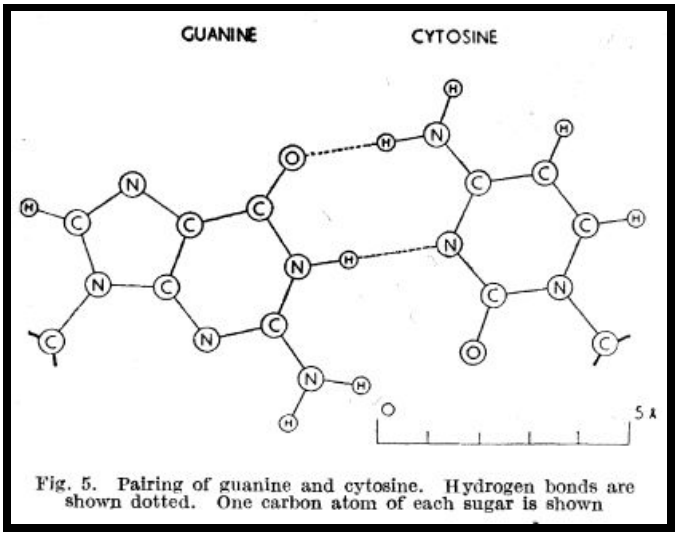

Fig. 5. Pairing of guanine and cytosine. Hydrogen bonds are shown dotted. One carbon atom of each sugar is shown

Figures 4 & 5 show base pairing. **First, to orient yourself, *highlight all of the hydrogen bonds in both figures*. *Write in where the sugar would be bonded to each of the nitrogenous bases.* It turns out Figure 5 has an omission. Can you find it? *Add it in once you've found it*. Need a hint? Look in our online notes for diagrams or in a class text.** Figures 4 & 5 are both drawn to scale. This means that the relative distances between atoms is accurate, but "blown up" so that we can see it. The scale bars in these figures are in angstroms, a unit used by scientists to measure distance on the molecular scale.

Measure the scale bar in Figure 4 in centimeters. Report the scale in the box at the right as "angstroms/centimeters."	
Determine the width of adenine in angstroms.	
Determine the width of thymine in angstroms.	
Determine the width of adenine + thymine in angstroms.	
Measure the scale bar in Figure 5 in centimeters. Report the scale in the box at the right as "angstroms/centimeters."	
Determine the width of guanine in angstroms.	
Determine the width of cytosine in angstroms.	
Determine the width of guanine + cytosine in angstroms.	

Did you get the same width for the (adenine + thymine) as the (guanine + cytosine)? What is the significance of this?

Define Purine and Pyrimidine.

Now moving on to the text of the paper, we will skip some paragraphs. However, if you would like to read the entire paper, it is linked on our website.

We have recently proposed a structure[1] for the salt of deoxyribonucleic acid which, if correct, immediately suggests a mechanism for its self-duplication. X-ray evidence obtained by the workers at King's College, London[2], and presented at the same time, gives qualitative support to our structure and is incompatible with all previously proposed structures[3]. Though the structure will not be completely proved until a more extensive comparison has been made with the X-ray data, we now feel sufficient confidence in its general correctness to discuss its genetical implications. In doing so we are assuming that fibres of the salt of deoxyribonucleic acid are not artefacts arising in the method of preparation, since it has been shown by Wilkins and his co-workers that similar X-ray patterns are obtained from both the isolated fibres and certain intact biological materials such as sperm head and bacteriophage particles[2,4].

Summarize this paragraph in 1 sentence:

The chemical formula of deoxyribonucleic acid is now well established. The molecule is a very long chain, the backbone of which consists of a regular alternation of sugar and phosphate groups, as shown in Fig. 1. To each sugar is attached a nitrogenous base, which can be of four different types. (We have considered 5-methyl cytosine to be equivalent to cytosine, since either can fit equally well into our structure.) Two of the possible bases—adenine and guanine—are purines, and the other two—thymine and cytosine—are pyrimidines. So far as is known, the sequence of bases along the chain is irregular. The monomer unit, consisting of phosphate, sugar and base, is known as a nucleotide.

This paragraph provides the reader with background knowledge that had been presented in their first paper. Highlight *five* pieces of background information that is presented in this paragraph.

The first feature of our structure which is of biological interest is that it consists not of one chain, but of two. These two chains are both coiled around a common fibre axis, as is shown diagrammatically in Fig. 2. It has often been assumed that since there was only one chain in the chemical formula there would only be one in the structural unit. However, the density, taken with the X-ray evidence[2], suggests very strongly that there are two.

Highlight the claim made in this paragraph.

What evidence do they cite?

The other biologically important feature is the manner in which the two chains are held together. This is done by hydrogen bonds between the bases, as shown schematically in Fig. 3. The bases are joined together in pairs, a single base from one chain being hydrogen-bonded to a single base from the other. The important point is that only certain pairs of bases will fit into the structure. One member of a pair must be a purine and the other a pyrimidine in order to bridge between the two chains. If a pair consisted of two purines, for example, there would not be room for it.

Highlight the claim made in this paragraph.

Use evidence from the measurements you made of Figures 4 & 5 to support the claim you highlighted.

We believe that the bases will be present almost entirely in their most probable tautomeric forms. If this is true, the conditions for forming hydrogen bonds are more restrictive, and the only pairs of bases possible are :

adenine with thymine ;
guanine with cytosine.

The way in which these are joined together is shown in Figs. 4 and 5. A given pair can be either way round. Adenine, for example, can occur on either chain ; but when it does, its partner on the other chain must always be thymine.

This pairing is strongly supported by the recent analytical results[5], which show that for all sources of deoxyribonucleic acid examined the amount of adenine is close to the amount of thymine, and the amount of guanine close to the amount of cytosine, although the cross-ratio (the ratio of adenine to guanine) can vary from one source to another. Indeed, if the sequence of bases on one chain is irregular, it is difficult to explain these analytical results except by the sort of pairing we have suggested.

Based on the information in these paragraphs, if a particular sample of DNA contains 35% adenine, how much of it will be composed of...

Guanine?

Cytosine?

Thymine?

The phosphate-sugar backbone of our model is completely regular, but any sequence of the pairs of bases can fit into the structure. It follows that in a long molecule many different permutations are possible, and it therefore seems likely that the precise sequence of the bases is the code which carries the genetical information. If the actual order of the bases on one of the pair of chains were given, one could write down the exact order of the bases on the other one, because of the specific pairing. Thus one chain is, as it were, the complement of the other, and it is this feature which suggests how the deoxyribonucleic acid molecule might duplicate itself.

How does this paragraph describe the author's proposed structure allowing for *variation*?

Define GENE.

Watson, who is still living in the US, went on to study zoology, molecular biology, and was involved in establishing the Human Genome Project.

Crick died in 2004, having worked in molecular biology, biophysics, and neurobiology until the day he died. He coined the term, "central dogma" that reflected the information flow from DNA □ RNA □ protein.

MUTATION is defined as "any change in DNA". We will find in our studies that these mistakes can result in varying levels of damage to organisms. These mistakes or changes in DNA can be inherited or caused by environmental factors. Nucleic acids and proteins are two different biochemical families. **In your own words, explain how a DNA mutation could affect the proteins produced by an organism.**