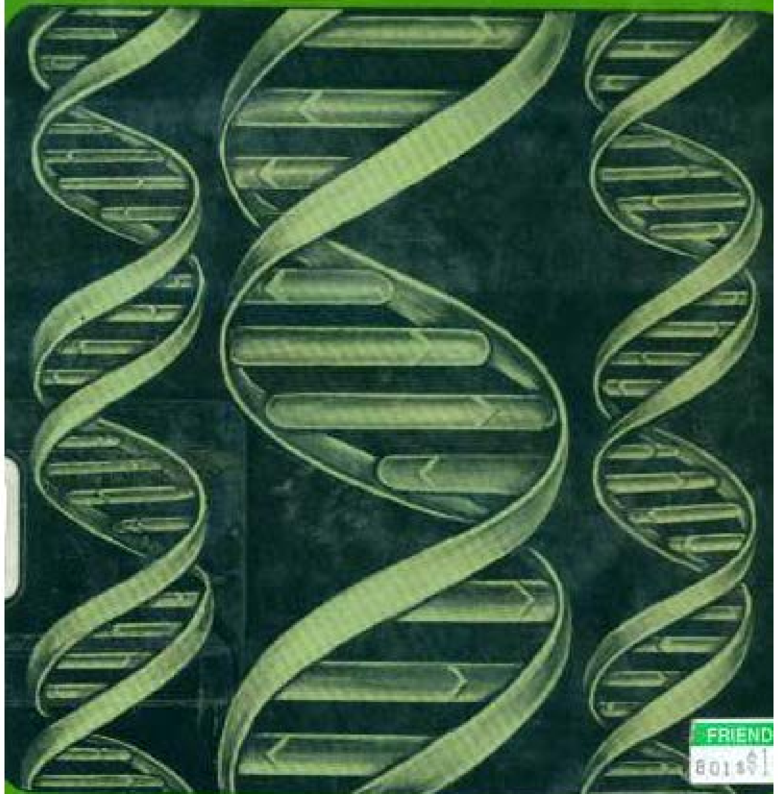
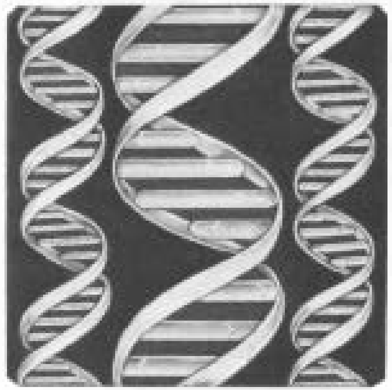
How Did We Find Out About DNA?

Isaac Asimov



illustrated by David Wool

How Did We Find Out About DNA?



Isaac Asimov

Illustrated by David Wool

Contents

1
The Pieces of Nucleic Acid

In 1869, a twenty-five-year-old Swiss chemist, Johann Friedrich Miescher (MEE-sher, 1844–1895), was working in the laboratories of a German chemist, Ernst Felix Hoppe-Seyler (HOH-puh-ZY-ler, 1825–1895).

Miescher was working with dead and broken-down cells. Cells are the tiny objects out of which the bodies of plants and animals are built.

In those days scientists were doing their best to find out about the substances that made up cells. Miescher was one of those working in this direction. He knew cells contained protein, which are very complicated substances, but he wanted to break them down into small pieces.

He added the enzyme (EN-zyme) pepsin (PEP-sin) to his material. An enzyme is a substance which acts to hasten certain chemical changes. Pepsin causes the large molecules of proteins to break down into small portions. But Miescher found there were other molecules in the cell that weren't touched by the pepsin.

Each cell has a nucleus, a small structure that is usually near the middle of the cell. The nucleus is enclosed by a thin membrane. Some of the molecules inside the nucleus remained unaffected by the pepsin.

Miescher separated this untouched material and tested it in certain chemical ways. He wanted to find what kind of atoms were present. Almost at once, he was surprised to find it contained atoms known as *phosphorus* (FOS-fuh-rus).

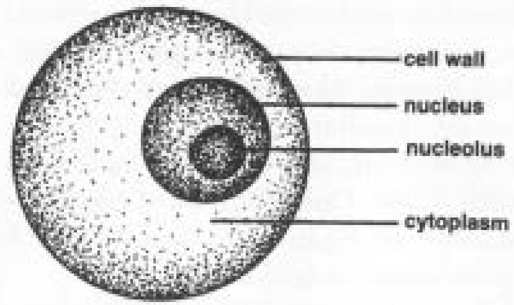
Phosphorus is not an unusual atom in general, but it was supposed to occur in rock. Until then, only one compound that contained phosphorus had ever been found in living tissue. It was a fatty substance named *lecithin* (LES-ih-thin) which had been discovered by Miescher's teacher, Hoppe-Seyler.

Miescher called the new material he discovered *nuclein* (NYOO-klee-in), because it was found inside the cell nucleus.

Miescher took his work to Hoppe-Seyler. The older chemist went over it and decided that Miescher ought not to announce the discovery just yet because he was young and inexperienced, and perhaps he had made a mistake. Hoppe-Seyler decided to go over all the work personally.

For two years Hoppe-Seyler worked most carefully, and finally he was satisfied when he found a very similar material in yeast cells.

The material he had obtained was a little different from that which Miescher had obtained, so they were given different names. Miescher's material could be easily obtained from an animal organ called the *thymus gland* (THY-mus), so it was named *thymus nuclein*. Hoppe-Seyler's material was from yeast so, of course, it was named *yeast nuclein*.



Cell

Another of Hoppe-Seyler's students was Albrecht Kossel (KOSS-ul, 1853–1927). In 1879, he began to study Miescher's nuclein.

Miescher had found that nuclein obtained from the sperm cells of salmon was attached to a very simple protein he called *protamine* (PROH-tuh-meen). He could separate them easily. Kossel decided to study this connection between nuclein and protein further.

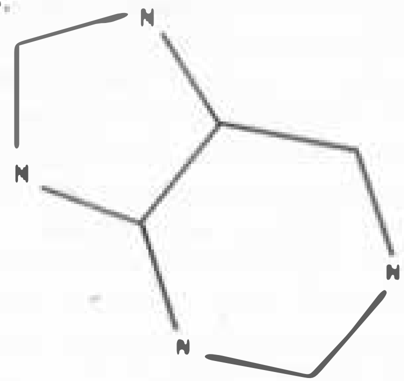
Kossel found that the nuclein he obtained was usually connected to a protein a bit more complicated than Miescher's protamine. Kossel called his protein *histone* (HIS-tone), from a Greek word meaning "cell." The combination of nuclein and protein is *nucleoprotein* (NYOO-klee-oh-PROH-tee-in).

He discovered that he could easily separate the histone from the nuclein. The reason they stuck together was that the nuclein behaved like an *acid* and the histone acted like a *base*. Acids and bases always interact with each other. Because of the acid behavior of nuclein, that material came to be called *nucleic acid* (nyoo-KLEE-ik-AS-id) and people began to speak of *thymus nucleic acid* and *yeast nucleic acid*.

At that time, no one had any notion what the molecules of nucleic acid were like, or how the atoms within those molecules were arranged. To find out, Kossel decided to treat the molecules chemically to break them up into smaller pieces.

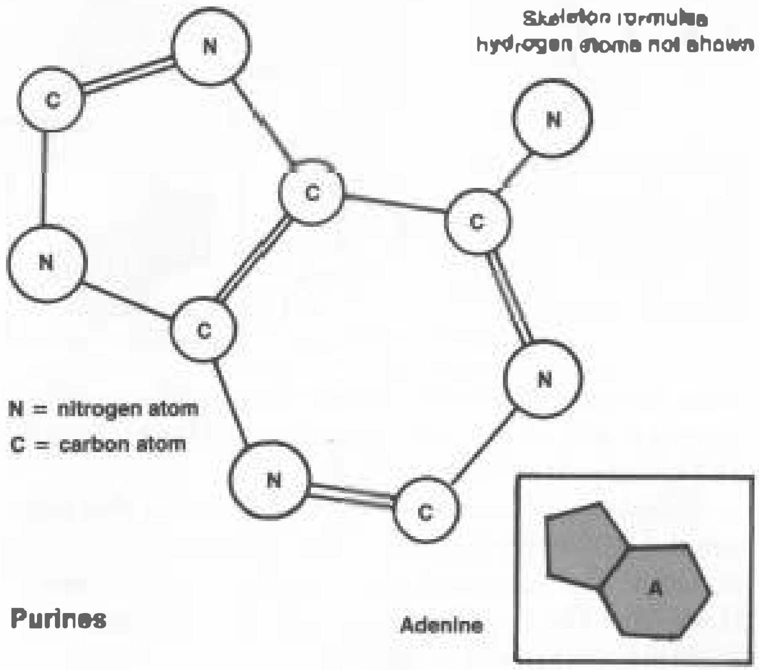
The smaller pieces might prove to be molecules that chemists already knew. Once they were recognized, it might be possible to figure out how to put them together to form a nucleic acid molecule.

Kossel and his students worked on the nucleic acids for years, and they managed to recognize some of the pieces. Some were made up of a double ring of atoms. The double ring consisted of a six-atom ring and a five atom ring that were joined so that two atoms were part of both rings.

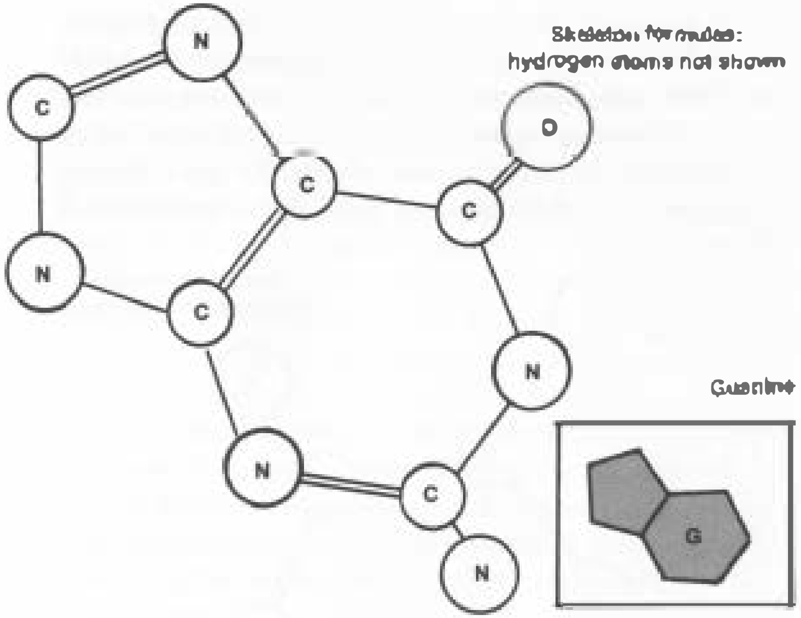


There is an atom at every angle of the double ring. If you count, you will see there are nine angles and, therefore, nine atoms. Four of the atoms are nitrogen, and they are marked by the Ns in the diagram. The other atoms are all carbon atoms.

A compound containing such a double ring in its molecule is known as a *purine* (PYOO-reen) to chemists. There are a number of such purines because the rings can have groups of additional atoms attached at one or more positions as *side-chains*. Every different side-chain or combination of side-chains results in a different purine.



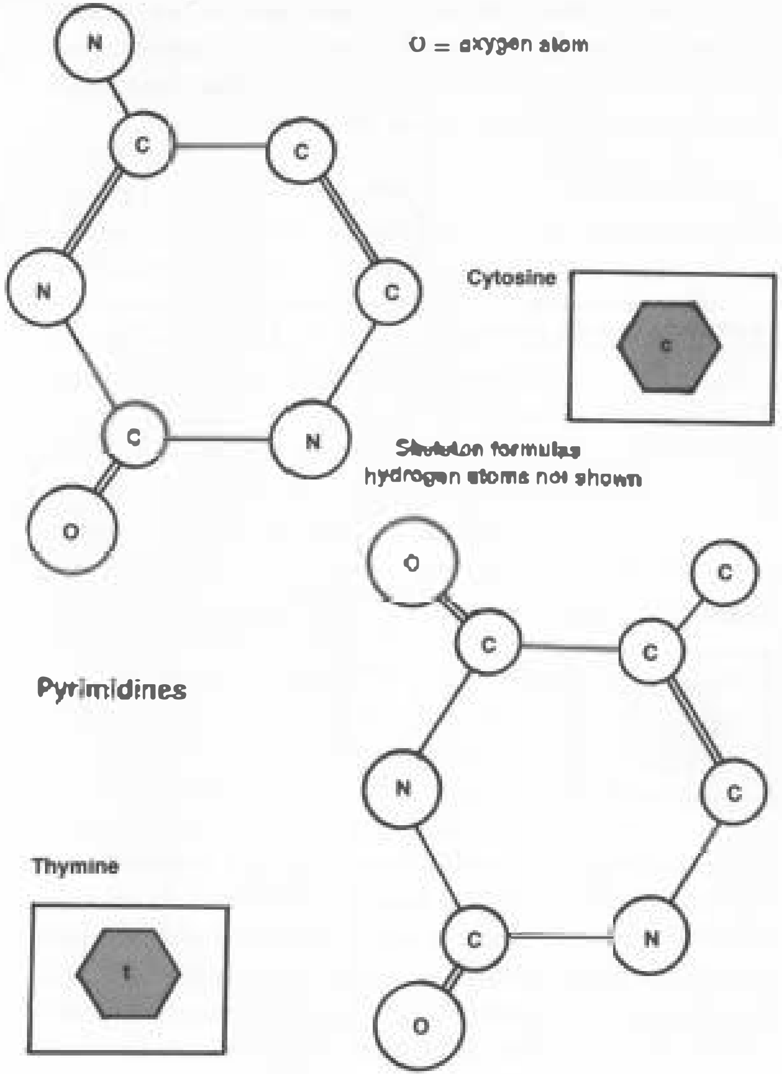
Chemists had studied some purines already. But Kossel found two purines that were new to chemists and that seemed to be part of every nucleic acid. One was *adenine* (AD-uh-neen) and the other was *guanine* (GWAH-neen). Adenine has an extra nitrogen atom attached and guanine has a nitrogen atom and an oxygen atom attached. These names are used very



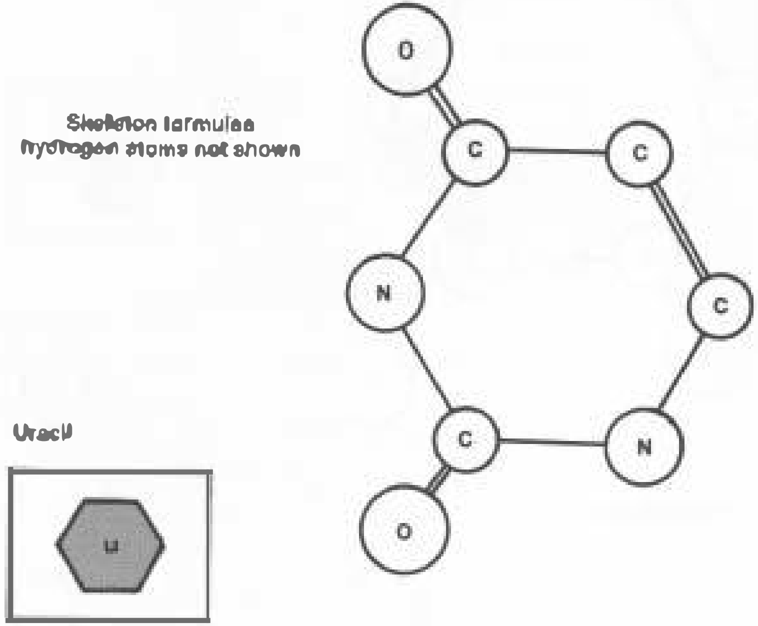
often in connection with nucleic acids, and sometimes just their initials are used. Adenine is referred to as A, and guanine is referred to as G.

Kossel also obtained pieces of nucleic acid that were simpler than the purines. This simpler kind had a molecule that contained only a single ring of six atoms. It was just like the six-atom ring in purines except the five-ring attachment is missing.

Such a ring is called a *pyrimidine* (pih-RIM-ih-deen) and there can be various pyrimidines since side-chains of different kinds can be attached to different positions on the ring.



Kossel found two pyrimidines among the pieces of thymus nucleic acid. One is *cytosine* (SY-toh-seen) and the other is *thymine* (THY-meen). Cytosine and thymine are also often referred to by their initials, *c* and *t*. I am using small letters in this case because the pyrimidines with their single ring have smaller molecules than the purines with their double rings and would seem to deserve small letters.



At last we find a difference in the molecules of thymus nucleic acid and yeast nucleic acid. Both have the two purines, adenine and guanine, and both have the pyrimidine, cytosine. Only thymus nucleic acid has thymine, however, which is why thymine has that name. Yeast nucleic acid has a different pyrimidine,

one that is very similar to thymine but not identical. This other pyrimidine is *uracil* (YOO-ruh-sil), and it can be represented by its initial, *u*. Thymine differs from uracil in that thymine has an extra carbon atom.

For his work on nucleic acids and for other work, too, Kossel was awarded the Nobel prize in physiology and medicine in 1910.

Of course, the purines and pyrimidines aren't all there is to nucleic acids. There were other pieces that Kossel had not identified. He thought that one of the additional pieces had the kind of structure that simple sugars have, but he wasn't sure.

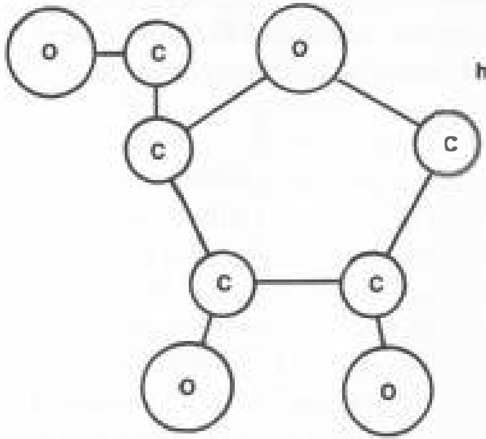
A Russian-American chemist, Phoebus Aaron Theodore Levene (1869-1940), traveled to Germany to study chemistry. One of the German chemists he studied with was Kossel, and he became interested in nucleic acids as a result. When he returned to the United States, he made them his life-work.

He broke down yeast nucleic acid molecules, and among the pieces that he obtained, he found the sugar molecule that Kossel had thought existed.

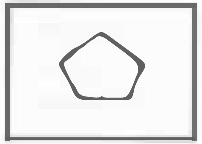
The simple sugars in living tissue that chemists knew about had six carbon atoms in their molecule, but the one that Levene had located had only five. Its molecule also contained ten hydrogen atoms and five oxygen atoms in addition to the carbon atoms.

Knowing just that wasn't enough because those atoms could be arranged into eight different but closely related sugars. Each one of these sugars had slightly different properties, and it was up to Levene to decide which of the eight varieties was the one he had obtained from yeast nucleic acid.

In 1909, Levene identified the sugar. It was one that chemists knew as *ribose* (RY-bose), or we can use the shortened form "rib."



Skeleton formulas
hydrogen atoms not shown

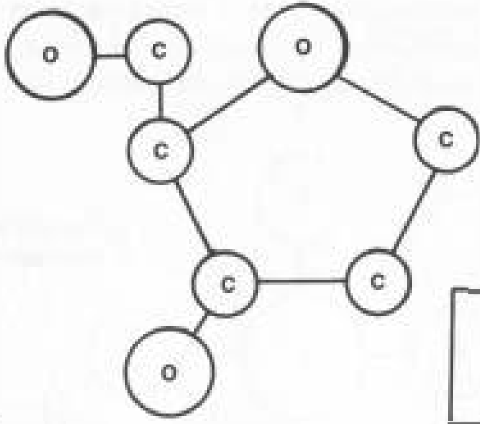


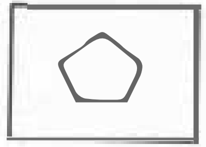
Ribose sugar

Levene had considerable trouble with thymus nucleic acid. It yielded a five-carbon sugar among its pieces. The five-carbon sugar of thymus nucleic acid, however, was not quite like any of the five-carbon sugars chemists knew about.

It was not until 1929 that Levene discovered what made this other five-carbon sugar different. It was exactly like ribose in its atomic arrangement except that one of the oxygen atoms was missing. Chemists had never worked with a sugar like that. Levene was the first ever to study such a molecule, so it's no wonder he had trouble.

Levene called the new sugar *deoxyribose* (dee-OK-see-RY-bose) where the "deoxy" was a Latin way of saying "minus an oxygen." We can use the abbreviation "derib" for it.



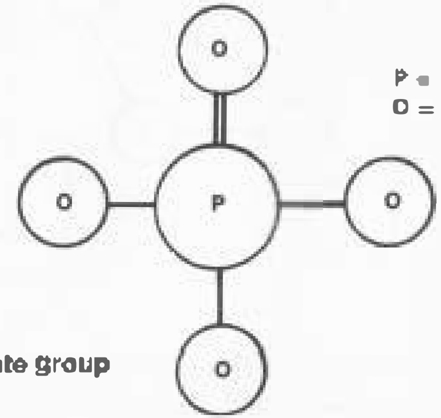


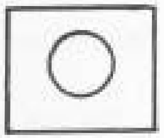
Deoxyribose

We can see now that, of the two varieties of nucleic acid, yeast nucleic acid had ribose and uracil in its molecules, while thymus nucleic acid had deoxyribose and thymine in its molecules.

Chemists have decided that the difference between ribose and deoxyribose is more important than the difference between uracil and thymine. They therefore took to calling yeast nucleic acid *ribonucleic acid* (RY-bo-nyoo-KLEE-ikAS-id) and thymus nucleic acid *deoxyribonucleic acid* (dee-OK-see-RY-boh-nyoo-KLEE-ik-AS-id). These two names are quite complicated, and they come up in speaking and writing so often that initials are commonly used. Ribonucleic acid is almost always referred to as *RNA* and deoxyribonucleic acid is almost always referred to as *DNA*.

Both RNA and DNA molecules also contain the phosphorus atoms that so astonished Miescher at the beginning. These phosphorus atoms do not occur in the nucleic acids by themselves. They are always part of a group containing oxygen and hydrogen atoms, too. The combination is called the *phosphate* group (FOS-fate), and we might use the abbreviation "ph" for it.





Phosphate group

Levene worked out what he thought was the way in which the various pieces fit together to make up a nucleic acid molecule. The purines and pyrimidines are attached to the ribose (or deoxyribose). The combination is attached to the phosphate group.

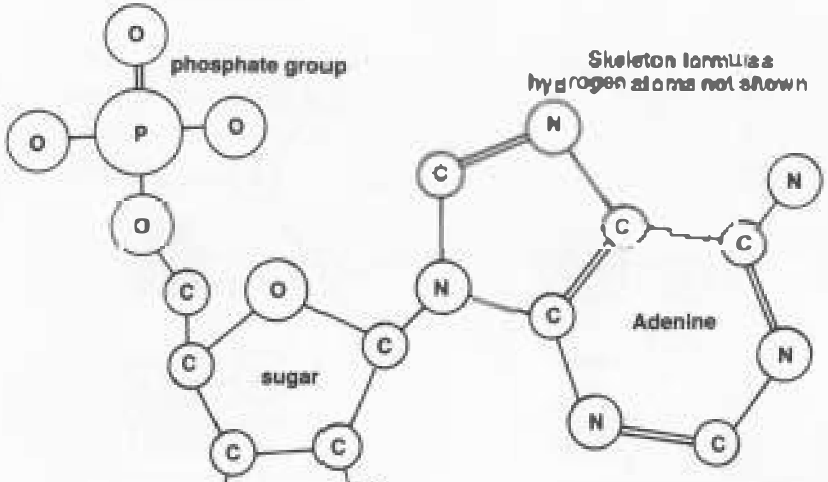
This is how one of the combinations would look in the RNA molecule:

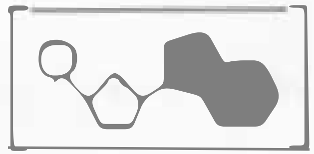
A – rib – ph

This is the combination in a DNA molecule:

A – derib – ph

In either case, such a combination is called a *nucleotide* (NYOO-klee-oh-tide).

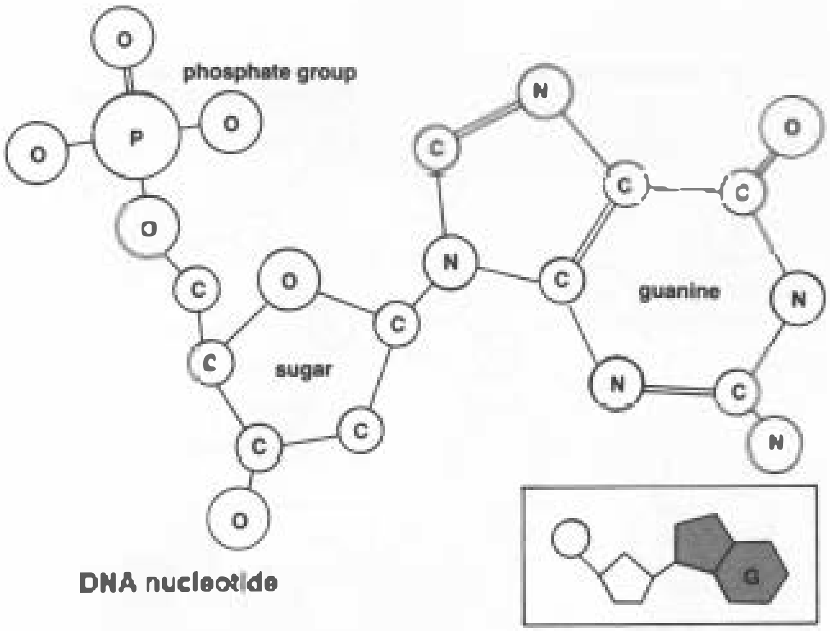




RNA nucleotide

There are four different nucleotides present in the RNA molecule, one of which has adenine (A) as part of the combination (shown above). The other three have guanine (G), cytosine (c), or uracil (u) as part of the combination. The DNA molecule also has four different nucleotides containing either adenine (A), guanine (G), cytosine (c), or thymine (t).

There are ways of measuring the total size of a molecule. Levene worked out the sizes of the nucleic acids he obtained from cells. It seemed to him that each molecule was large enough to be made up of four nucleotides, very likely one of each kind. The nucleotides held together because the phosphate group of one nucleotide made a second connection with the ribose (or deoxyribose) of the next nucleotide.



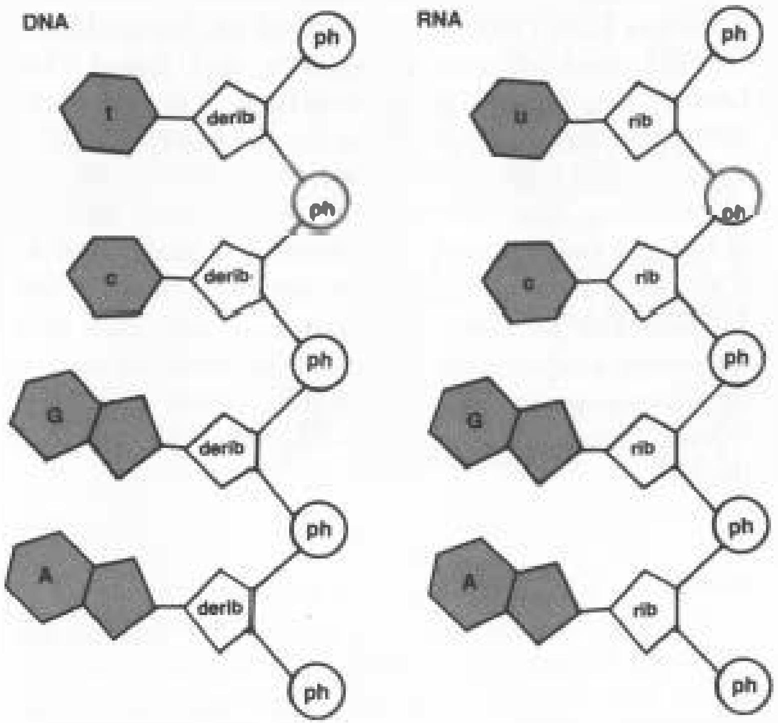
DNA nucleotide

The four nucleotides, clinging together, are a *tetranucleotide* (TET-ruh-NYOO-klee-oh-tide), where the "tetra" is from the Greek word for "four."

This is the way a DNA tetranucleotide and an RNA tetranucleotide might look:

DNA		RNA	
A – derib		A – rib	
	ph		ph
G – derib		G – rib	
	ph		ph
c – derib		c – rib	
	ph		ph
t – derib		u – rib	

Since the *derib-ph* and the *rib-ph* are always the same from nucleotide to nucleotide, a simpler way of showing what the tetranucleotides might look like would be as follows:



Tetranucleotides of DNA and RNA

The best way of making certain that Levene was correct about the structure of the nucleotides was to start with simple molecules. These simple molecules would be treated chemically in a way designed to put them together in a known fashion. Finally the structure Levene had reasoned out could be put together. The properties of the built-up structure would then be studied. If they proved to be the same as those of the nucleotides obtained from nucleic acids, then Levene would be proved right.

Beginning in 1938, a Scottish chemist, Alexander Robertus Todd (1907–), worked on this problem. He prepared all the nucleotides and found that Levene was correct in his structure. For his work, Todd received the Nobel prize for chemistry in 1957.

(You might wonder why Todd got a Nobel prize just for showing that Levene was right, when Levene never got one himself. The answer is that, even in science, things aren't always perfectly fair. When Levene did his work, no one had any idea how important nucleic acids were. By the time the importance was recognized and Nobel prizes were frequently given out for work with them, Levene had died.)

2 Nucleic Acids? Proteins?

BIOCHEMISTS WONDERED WHAT nucleic acids do in the body. Do they have an important part to play?

It was possible they might. Miescher, in the early days of his discovery, had found nucleic acid in the sperm cells of fish. Sperm cells are very tiny objects that don't have room in them for anything except the father's genes,* which carry inherited characteristics. A sperm cell enters an egg cell that carries the mother's genes. The *fertilized egg cell* that results develops into a new organism. Could nucleic acids therefore have something to do with inheritance? If so, they might be very important indeed.

In 1914, a German biochemist, Robert Joachim Feulgen (FOIL-gen, 1884–1955), found a red dye that would combine with DNA, but not with RNA. In

*It would be helpful in reading this book, if you also read my book *How Did We Find Out About Genes?* (Walker, 1983)

1923, he tried this dye on thin slices of living cells. The dye poisoned the cells, but it combined with material in some parts of the cell while leaving other parts untouched. Wherever it combined, DNA had to be present, and there would be a deep red stain in that spot, while everything else remained colorless. It was like making a colored map of the cells, showing where the DNA was.

It turned out that DNA was located almost entirely in the nucleus of every plant and animal cell tested.

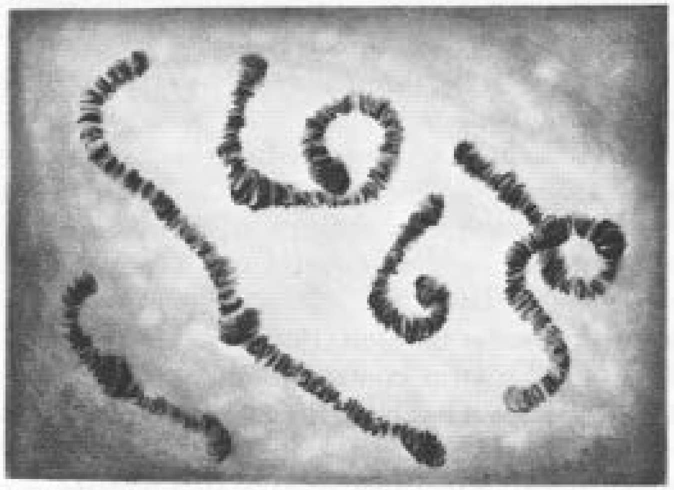
In the 1940s, a Swedish biochemist, Torbjorn Oskar Caspersson (1910-) went further. There are some enzymes that break up DNA but leave RNA untouched. There are other enzymes that break up RNA and leave DNA untouched. Caspersson used each enzyme on different cells, and thus produced some cells that had only DNA and other cells that had only RNA. He then shone ultraviolet light through the cells. Ultraviolet light is absorbed in a particular way by either kind of nucleic acid. Caspersson could tell exactly where both DNA and RNA were in the cells.

He found the DNA in nuclei was actually located in the *chromosomes* (KROH-moh-somez). The RNA, on the other hand, was located outside the nucleus in the *cytoplasm* (SY-toh-plaz-um).

With further investigation, some RNA was found in parts of the nucleus and some DNA in parts of the cytoplasm. However, almost all of the DNA was in the chromosomes, and most of the RNA was in the cytoplasm.

By that time, scientists knew very well that the chromosomes in the nuclei—tiny objects that looked like stubby bits of spaghetti—were deeply involved

with inheritance. The chromosomes carried the genes, and it was the genes passing from parents to children that carried all the characteristics of that particular organism.



Chromosomes of the fruit fly

With DNA located in the chromosomes, it therefore looked as though DNA might have something to do with inheritance.

Of course, not all living things are composed of cells. There are tiny objects called *viruses* that are far smaller than cells, and that seem to be able to get inside cells and multiply there. Such viruses, in multiplying, produce other viruses just like themselves, so they must have some device for passing on their characteristics by inheritance. What would that device be?

Until biochemists managed to get pure samples of viruses without any pieces of cells included, they couldn't be sure what made up the viruses. Pure samples were first obtained by an American biochemist, Wendell Meredith Stanley (1904–1971). He was studying the *tobacco mosaic virus* that caused a disease in tobacco plants. In 1935, he managed to get fine, needlelike crystals out of mashed-up tobacco leaves that were infected with the virus.

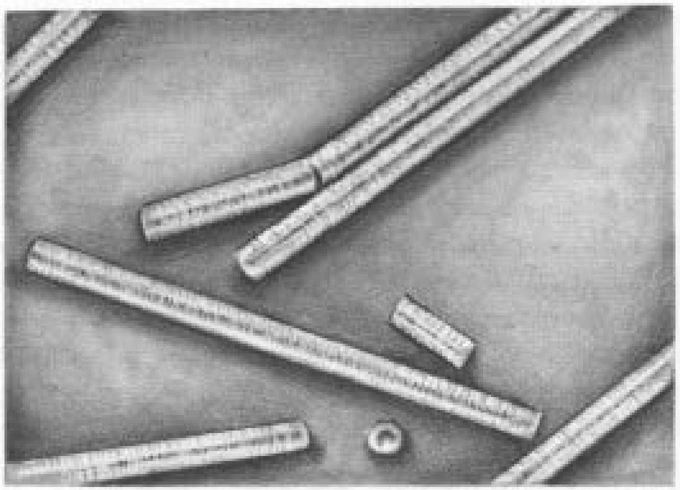
The crystals, which were pure tobacco mosaic virus, were made up of proteins. Since then, all viruses that have been isolated and purified have been found to be made up of protein. For his work, Stanley won a share of the Nobel prize for chemistry in 1946.

Almost at once it was found that viruses contain more than protein. In 1937, an English biologist, Frederick Charles Bawden (1908–) found that the tobacco mosaic virus contained RNA as well as protein. Since then it has been discovered that all viruses also contain nucleic acid. The simpler viruses usually contain RNA, but more complicated viruses have DNA. Some have both.

It is possible to think of viruses as being something like loose chromosomes that are not part of cells. When a virus invades a cell, it somehow seizes control of the cell from the cell's own chromosomes.

What is there about chromosomes and viruses that controls inheritance, as well as the day-to-day workings of cells and organisms? Since all chromosomes and viruses are made up of proteins and DNA (except for the simpler viruses which have RNA), it might be the protein, the DNA, or both.

At first scientists felt sure that it must be the protein molecules that controlled living cells and were responsible for inheritance. They thought that whatever it



Tobacco mosaic virus

was that DNA did, it must serve only as an assistant to the protein.

One reason scientists felt sure of this was that for a whole century they had looked upon proteins as the most complicated molecules involved with life. Indeed, they seemed the most complicated molecules that existed anywhere.

Proteins are made up of giant molecules, each one being made up of anywhere from hundreds to hundreds of thousands of atoms. There are other giant molecules too, such as those of starch, or of the cellulose that makes up a major portion of wood. It seemed no other giant molecules, however, could be compared to those of proteins.

Giant molecules can be easily broken up into small units. Such small units are hooked together like beads on a string to form the giant molecules. Usually the units that make up giant molecules are only one kind. Thus, starch can be broken down into units that are molecules of a simple sugar called glucose. Cellulose can also be broken down into glucose units.

Plastics are made up of giant molecules that chemists produce in the laboratory. These can be broken down into simple units that are of only one or possibly two varieties.

Protein molecules can also be broken down into simple units, and these are all *amino acids* (uh-MEE-noh-AS-idz). Amino acids come in many varieties, however. Protein molecules consist of long strings of as many as twenty different types of amino acids.

Molecules of starch, cellulose, or plastics, built up of only one or two different units, can differ from each other only in the length of the chain, or in whether the chain is straight or branched.

Chains of amino acids, however, in forming protein molecules, can differ not only in the number of each of the twenty different units, but also in the exact arrangement. There are trillions of different arrangements possible, and each of these represents a different molecule.

This means that every species—every individual living thing—can have proteins not quite like those of any other species or individual. The reason living organisms vary among themselves and can do so many complicated and different things must be because they are made up of proteins that can take on many shapes and have many different properties.

According to Levene's findings, nucleic acids didn't consist of giant molecules. They were made up of only four different nucleotides and, in the nucleic acid

molecule, there seemed to be only one each of these four nucleotides.

The workings of a cell depend on the presence of a large number of different enzymes, each in a particular quantity. These enzymes are made up of protein molecules and, since each new cell can make the proper enzymes for itself, somewhere in the cell there must be the information (or "blueprint") needed for enzyme molecules to be built up correctly.

It seemed certain that only a protein molecule could be complicated enough to carry the blueprint of another protein molecule. The little four-nucleotide molecules of DNA seemed to be too small to perform the task.

To be sure, Levene turned out to be wrong in supposing that nucleic acid molecules had only four nucleotides apiece. The methods he used to extract the nucleic acids from the cell were pretty violent, and they reduced the molecules to small pieces.

As biochemists learned to extract the nucleic acid molecules by gentler methods, they found they ended up with larger and larger molecules. Eventually, it began to seem that DNA also consisted of giant molecules, as large or larger than protein molecules.

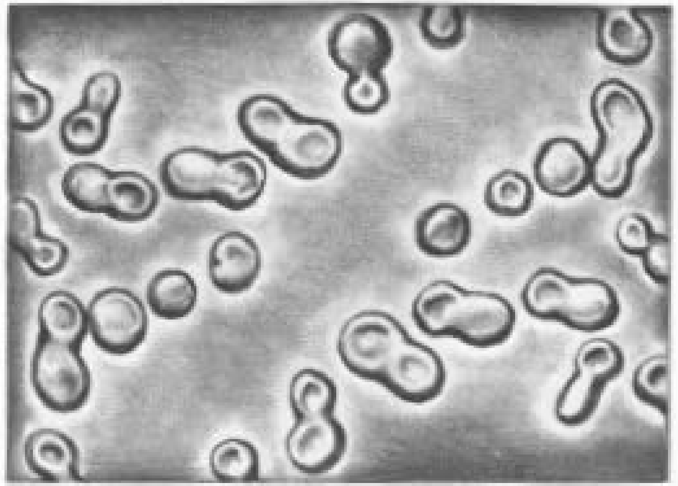
Yet biochemists were so used to thinking of proteins as the most important molecules in living cells that they continued to ignore DNA. And then one day, everything changed.

3
The Winner: DNA

FOR A LONG time, scientists had been studying pneumococci (NYOO-moh-KOK-sigh), the bacteria that give rise to the lung disease of pneumonia. A single such bacterium is a *pneumococcus* (NYOO-moh-KOK-us).

Pneumococci came in two varieties. In one, each bacterial cell is covered with a smooth coat made of complicated sugarlike molecules. This variety is *pneumococcus S* (for "smooth"). The other variety of pneumococcus lacked the smooth coat, and therefore had a rough appearance. It is *pneumococcus R* (for "rough").

Apparently, pneumococcus R lacks a gene that is needed to help produce the complex material that forms the coat. In 1928, an American biologist, Fred-erick Reece Griffith, Jr. (1891–), heated a quantity of pneumococcus S until all the bacteria were killed. He then added the liquid containing the dead bacteria to a quantity of living pneumococcus R. He found that pneumococcus R, in multiplying, became pneumococcus S.



Pneumococcus bacteria

The pneumococcus S, although dead, still contained the gene that produced a smooth coat, and that gene remained in working condition. When the gene was added to the pneumococcus R, which lacked such a gene, that bacterium could form the coat so that it then became pneumococcus S.

Naturally, scientists tried to isolate the gene, or the *transforming principle*, as it was called. They all assumed that the gene would prove to be some sort of protein.

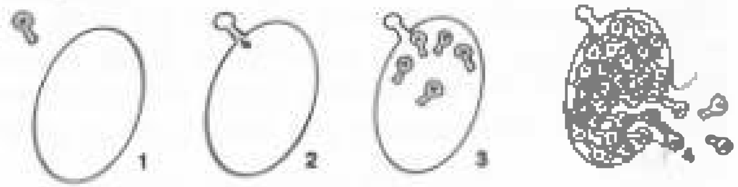
A Canadian-American biologist, Oswald Theodore Avery (1877–1955), was particularly interested in the problem. Little by little he purified the fluid containing the transforming principle until he had gotten rid of everything but the transforming principle itself.

When he tested the solution to see what it contained, he found DNA. He could find no protein in it. In 1944, Avery, along with two co-workers, announced that DNA, without any help from protein, was the gene.

If this were so, it was very likely that all genes were made up of DNA, and that DNA was the material in cells that controlled their behavior. It was DNA that carried the blueprint for inherited characteristics and passed these on from cell to cell, when cells divided, and from parents to their young, when plants and animals reproduced themselves.

Once Avery made his discovery, scientists found more and more reason to suppose that DNA controlled the cell. In 1952, for instance, an American biologist, Alfred Day Hershey (1908–), found that when a virus invaded a cell, only the DNA of the virus entered the cell. The protein of the virus remained outside. Yet inside the cell, not only did the virus-DNA form more DNA exactly like itself, but it also formed a great deal of protein like the protein that had been left outside. This made it clear that DNA molecules carried the blueprint for protein molecules.

Now scientists had to find what there was about the DNA molecule that made up the blueprint, and how it could multiply itself and still keep the blueprint intact.



Virus attacking and destroying a cell

In 1944, just as Avery was announcing his finding about DNA, a new way of analyzing complicated mixtures of substances was worked out. It was called *paper chromatography*, a process in which a chemical mixture is separated into its components. The new technique was quickly used to investigate DNA.

An Austrian-American biochemist, Erwin Chargaff (1905–), broke up the DNA molecule until all the purines and pyrimidines were loose. He then analyzed the mixture of the two purines (adenine and guanine) and the two pyrimidines (cytosine and thymine) to see how much there was of each one.

In 1948, Chargaff showed that in all the DNA he had tested, the total number of purine molecules was always equal to the total number of pyrimidine molecules. This meant that adenine plus guanine was always equal to cytosine and thymine (A + G = c + t). What's more, the number of adenine molecules was always equal to the number of thymine molecules (A = t), and the number of guanine molecules was always equal to the number of cytosine molecules (C = c).

Chargaff couldn't tell why this should be, but these equalities had to have something to do with the structure of the DNA molecule.

Another way of checking on the structure of the DNA molecule was to purify a solution of it until only DNA molecules were present and then get them out of the solution in the form of delicate fibers. Those fibers could be bombarded with X rays, that will strike one atom in the molecule and bounce off in some direction.

Suppose a very long molecule, such as DNA, has a

repeated structure. Also suppose there is a particular group of atoms in one place and that it appears again and again at regular intervals along the chain (like the design that keeps repeating, every foot or so, on wallpaper.)

In that case, the X rays hit the group of atoms at every place it appears in the molecule and bounce off, each time in the same direction. Instead of X rays bouncing every which way, you have them forming a sizable beam, all in the same direction.

X rays can cloud a photographic film, so that you can take a picture of them. If molecules contain groups of atoms without any particular pattern, the X rays bounce off in all directions and the picture just shows a foggy haze.

If, however, the molecules have a repeated structure, x-ray beams come off only in certain directions, and the picture shows a pattern of dots. From the pattern of dots in such an *x-ray diffraction* picture, it might be possible to deduce the kind of repetitions there are in a molecule and to get a notion of its structure. It might become possible to build a three-dimensional model of the molecule, with all the atoms in the right place.

In 1951, an American chemist, Linus Carl Pauling (1901–), worked on protein structure, making use of x-ray diffraction and of his own studies of the way in which atoms fit together. He showed that chains of amino acids would twist into the shape of a *helix* (HEE-liks). Familiar objects that have the shape of a helix are spiral staircases and bed springs.

Of course, only very simple proteins have their amino acid chains twisted into long helixes in a

Linus Pauling

straight line. The helixes bend and curve in complicated ways in proteins such as enzymes. Still, once biochemists began to think about helixes, the further complications could be worked out. The exact shape of many protein molecules is now known.

When Pauling made his suggestion, biochemists were struck by it, and some thought that the long chains of nucleotides that make up the DNA molecules were also twisted into helixes.

4 The Double Helix

TWO SCIENTISTS WHO were particularly interested in the possibility of DNA helixes were an Englishman, Francis Harry Compton Crick (1916–) and his American co-worker, James Dewey Watson (1928–).

They tried different types of helixes, but none of them seemed suitable. A helix had to twist according to the natural way in which the atoms fit together. It had to be the kind of helix that would explain the x-ray diffraction photos. It also had to help explain how nucleic acid worked. No matter what Crick and Watson tried, nothing seemed to fit all of the requirements.

What they needed were good x-ray diffraction photos of pure DNA, but those weren't easy to get.

As it happened, at the place where Crick and Watson were working, there was a New Zealand-born biochemist, Maurice Hugh Frederick Wilkins (1916–). He prepared pure DNA fibers that would be expected to produce particularly good x-ray diffraction photos. Working for him was an English chemist, Rosalind Elsie Franklin (1920-1958), who used Wilkins' DNA fibers to make the very best x-ray diffraction photos that had yet been taken.

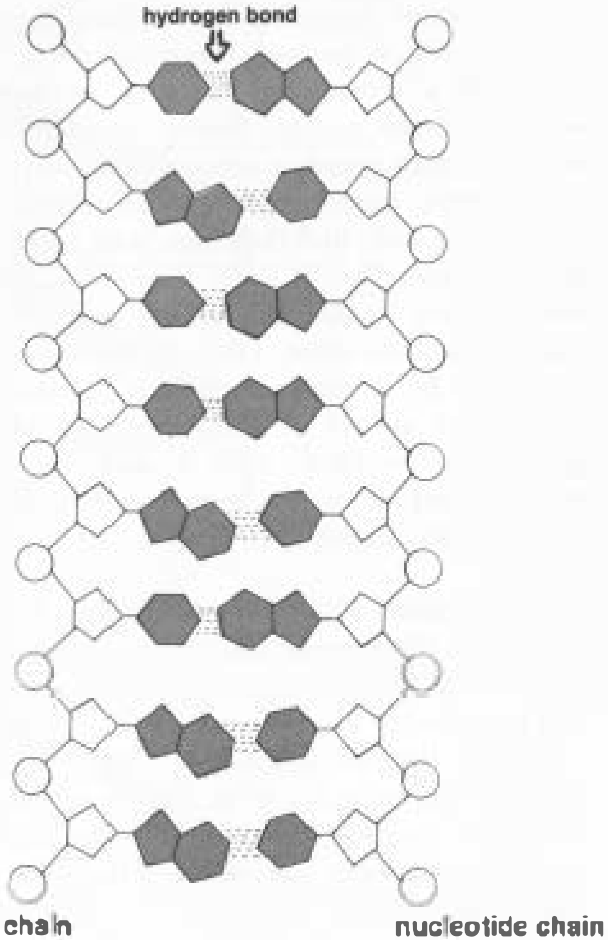
Rosalind Franklin

Franklin was a very cautious scientist and did not wish to hurry in figuring out the meaning of her photos. She did not want to make any mistakes nor did she want anyone else to see them while she was thinking out the meaning.

Wilkins, however, showed her photos to Watson and Crick without asking her permission. They (Watson, especially) were much less cautious than Franklin, and the photos gave them a remarkable idea almost at once.

Watson and Crick decided that whereas a protein molecule was made up of a chain of amino acids, a DNA molecule was made up of a *double* chain of nucleotides. The two chains of nucleotides were so arranged that the purine and pyrimidine portions faced each other.

The purines and pyrimidines of the two nucleotide chains are held together by what chemists call *hydrogen bonds*. The hydrogen bond is much weaker than the ordinary bonds that hold atoms together in a molecule. It is strong enough to hold the two nucleotide chains together under ordinary circumstances, but at crucial moments, the two chains can be pulled apart.



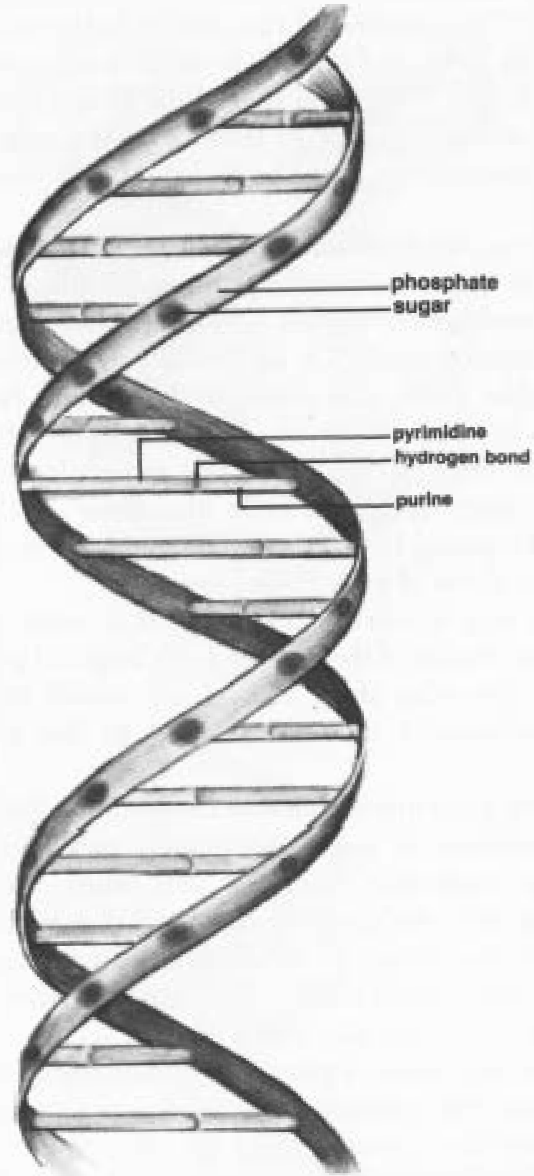
The two chains are, in fact, like the opposite teeth of a zipper. The zipper usually holds together under the proper conditions, but when you pull down the slide, the two sides of the zipper pull apart easily.

In order to even the space between the nucleotide chains, a double-ring purine on one side must always be opposite a single-ring pyrimidine. This makes the space between the nucleotide chains just wide enough to hold three rings together at every position. If a purine faced a purine, there wouldn't be enough room. If a pyrimidine faced a pyrimidine, they wouldn't reach across. Either way, the two chains wouldn't cling together properly at those points.

Hydrogen bonds hold together adenine and thymine quite well, and they also hold together guanine and cytosine. The fit is far worse and the hydrogen bonds much weaker in adenine and cytosine, or in guanine and thymine. On one nucleotide chain the order can be anything, but then, on the other side, there must always be a matching set of purines or pyrimidines—c for G, t for A, and vice versa. If the arrangement on one nucleotide is A, G, A, t, t, c, G, C, G, c; then on the other it must be t, c, t, A, A, G, c, c, c, C.

This explains why Chargaff had found that in DNA molecules there were always equal numbers of adenines and thymines and equal numbers of guanines and cytosines.

Finally, the two chains were twisted in such a way as to produce a *double helix*. They were like two spiral staircases, twisting together, with the banister of one just fitting between the curves of the other.



Double Helix of DNA

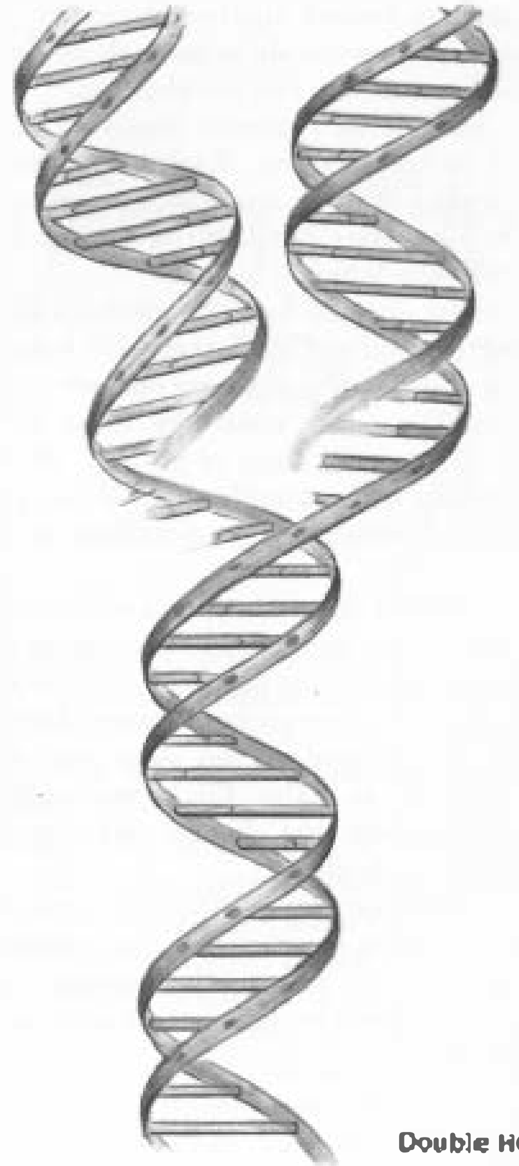
Watson and Crick described this double helix structure of DNA in 1953, and it made a sensation at once. Watson, Crick, and Wilkins all shared the Nobel prize for medicine and physiology in 1962. Franklin might have been honored, too, but she had died four years earlier.

The structure worked out by Watson and Crick explained how DNA molecules produced replicas of themselves (*replication*, rep-lih-KAY-shun) when cells divided. Because each DNA molecule can produce another just like itself, skin cells can divide into two skin cells, liver cells can divide into two liver cells, and so on. It is why egg cells have DNA molecules like those of the mother, sperm cells like those of the father, and the young have DNA molecules and characteristics like those of both their parents.

Here's the way it works. When a cell is ready to divide, the two chains of the double helix begin to pull apart. The nucleotides along each chain quickly pick up single nucleotides that are present in the cell fluids.

Thus, when a thymine-adenine combination pulls apart, the thymine on one chain quickly picks up a new adenine nucleotide from the cell fluids. That would be the only molecule in the fluid that would stick neatly to the thymine. Meanwhile, the adenine on the other chain would pick up a thymine nucleotide which would, of course, also give a good fit.

In the same way, when a guanine-cytosine combination pulls apart, the guanine picks up a new cytosine nucleotide, and the cytosine picks up a new guanine nucleotide. The new nucleotides, as they cling to the chain, form a chain of their own. The result is that, by



Double Helix replicating

the time a double helix has pulled entirely apart, each nucleotide chain has formed another chain that fits itself. There are now two double helixes where there was only one before, and the two are identical.

In this way, each DNA molecule forms an exact replica of itself at cell division. Each chromosome forms an exact replica of itself, and each new cell gets a new set of chromosomes just like the set that had existed in the cell from which it formed.

A young organism, formed from the mother's DNA molecules in the egg cell and the father's DNA molecules in the sperm cell, will have some characteristics like those of its father and some like those of its mother. If the DNA replication is perfect, all the baby's characteristics will resemble those of one parent or the other, or sometimes be something in between.

But nothing is perfect. Every once in a while something goes wrong in the replication. The wrong nucleotide may happen to get into position and be locked in place before it can bounce away. In some places, a nucleotide may be skipped, or an extra one may accidentally squeeze in. The DNA molecule is changed as a result and that change may then be preserved in future replications.

The result is a *mutation* (myoo-TAY-shun) in which a new cell, or a new living creature, has characteristics that are not found in the cell, or the parents, from which it arose. It is these mutations that help make evolution proceed.

5 From Triplets to Amino Acids

So FAR, the process of replication only explains how DNA molecules produce other DNA molecules. How do they control cells and living creatures? They do so by guiding the production of enzymes which control chemical changes in cells. The order of the nucleotides along the molecule of DNA must control the order of amino acids along the molecule of an enzyme.

But how can that be? Each enzyme molecule has a particular arrangement of twenty different enzymes. DNA, on the other hand, is made up of some particular arrangement of only four different nucleotides. How can information about twenty different amino acids be squeezed into four different nucleotides?

It's not impossible. English words are made up of twenty-six different letters of the alphabet, yet any

English message can be transmitted by means of the Morse code, which contains only two different signals—dots and dashes. The reason is that a combination of *several* dots and dashes can be used to stand for a particular letter of the alphabet. Out of just dots and dashes, many combinations can be made.

In 1954, soon after Watson and Crick had worked out the structure of DNA, the Russian-American scientist, George Gamow (GAY-mov, 1904-1968), suggested that it was a group of nucleotides, and not just one only, that stood for a particular amino acid.

For instance, suppose you moved along the nucleotide chain of DNA studying every pair of nucleotides. The first nucleotide of a pair could be any one of the four nucleotides and the second nucleotide could be any one of the four, also. That means the total number of different pairs is 4 × 4, or 16. You can write down all the combinations such as GG, GA, GC, CG, AC, AT, and so on, and when you've written down every possible one, you will have sixteen.

Of course, sixteen pairs is not enough, but suppose you take the nucleotides three at a time. Then you would have 4 × 4 × 4, or 64 different combinations possible. That's more than enough. It would be possible to have each of two, three, or even four different *triplets* all standing for the same amino acid. You could have a triplet that marks off the beginning of an amino acid chain, and another that marks off the end.

But how does the information contained in the DNA molecule—the order of its nucleotide triplets—get to the place in the cell where enzyme molecules are manufactured. The DNA is in the chromosomes that are inside the cell nucleus. The enzymes, however, are manufactured in the cytoplasm that is outside the cell nucleus.

The place of manufacture of the enzymes was discovered, in 1956, by the Rumanian-American biologist, George Emil Palade (pah-LAH-dee, 1912–). He made use of an *electron microscope* which could magnify the interior of cells much more than ordinary microscopes could. He found tiny structures in the cytoplasm, perhaps 150,000 of them in each human cell, where the enzymes were manufactured. Each of these structures contained a great deal of RNA (ribonucleic acid) so they came to be called *ribosomes* (RY-boh-somez).

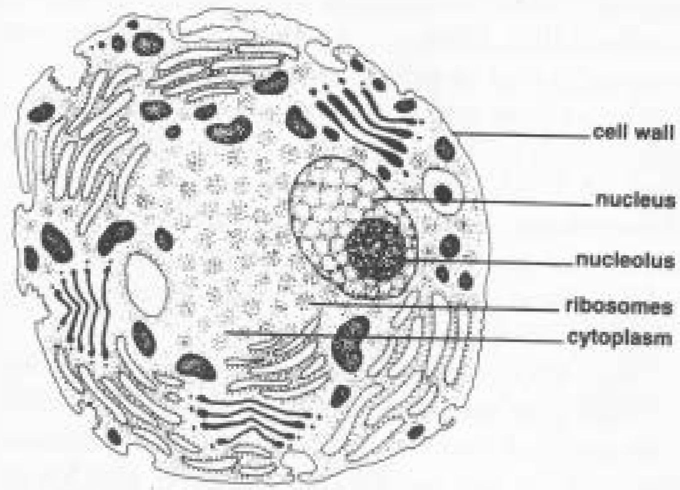


Diagram of a typical cell

For this work, Palade got a share of the 1974 Nobel prize for physiology and medicine.

But how did the information on the DNA molecule get from the chromosomes to the ribosomes?

Nobel Laureates

François Jacob
1965

Jacques Monod
1965

In 1961, two French biochemists, Jacques Lucien Monod (moh-NOH, 1910–1976) and François Jacob (zhah-KOBE, 1920–) suggested that RNA was the answer to that question. After all, RNA is found both in the nucleus and in the cytoplasm (in the ribosomes particularly). Then, too, RNA has a structure just like that of DNA, except that it has ribose instead of deoxyribose, and uracil instead of thymine. When a DNA molecule replicates, it might, every once in a while, form an RNA nucleotide chain, instead of another DNA nucleotide chain.

The RNA nucleotide chain would have its nucleotides in the exact order of the DNA nucleotide chain except that there would be a uracil (u) where a thymine (t) had been. The molecule of RNA would then slip out of the nucleus and would serve as the messenger carrying the DNA information to the ribosomes. This molecule of RNA came to be known as messenger-RNA.

Monod and Jacob proved to be correct, and they each received a share of the 1965 Nobel prize for

medicine and physiology for this and for other work on nucleic acids.

Meanwhile, in 1955, a Spanish-American biochemist, Severo Ochoa (oh-CHOH-ah, 1905-), had discovered an enzyme that acted to tie nucleotides together to form an RNA chain. This made it possible to make synthetic RNA. Ochoa received a share of the 1959 Nobel Prize for medicine and physiology.

Once the matter of messenger-RNA had been worked out, an American biochemist, Marshall Warren Nirenberg (1927–), manufactured synthetic messenger-RNA. By choosing which nucleotides to start with, he could make messenger-RNA with particular triplets. He would then find out what particular amino acids would be produced in an amino acid chain. In this way, Nirenberg began to break the *genetic code*. He discovered which nucleotide triplet produced which amino acid. By 1967, every single

Nobel Laureates

Albrech Kossel
1910

Wendell Stanley
1946

Alexander Todd
1957

nucleotide triplet was tied to a particular amino acid, and the genetic code was completely worked out.

Nirenberg, and an Indian-American chemist, Har Gobind Khorana (1922-), who worked with him, received shares of the 1968 Nobel prize for medicine and physiology.

While Nirenberg was working on the genetic code, another American biochemist, Mahlon Bush Hoagland (1921-) was locating small RNA molecules in the cytoplasm.

These were double-ended molecules. At one end was a nucleotide triplet that would fit a particular triplet on the messenger-RNA. On the other end was a portion that attached itself to a particular amino acid. Because they transferred the information from the triplet to the amino acid, such molecules were called *transfer-RNA*.

You can see now how it works. The messenger-RNA is formed with the copy of a section of the DNA

Nobel Laureates

Max Delbrück
1969

Salvador Luria
1969

molecule. The messenger-RNA travels to the ribosome where transfer-RNA molecules attach themselves to the various triplets of the messenger-RNA. Each triplet gets the transfer-RNA molecule that happens to fit that particular triplet and no other. To the other end of each transfer-RNA, amino acids attract themselves, just the amino acid that particularly fits that end of that particular transfer-RNA. The amino acids then all hook together, and you have a particular enzyme molecule.

The transfer-RNAs were studied carefully by an American chemist, Robert William Holley (1922-). He purified several varieties in 1962. By 1965, he actually put the proper nucleotides together and synthesized one of them. In 1968, he shared the Nobel Prize with Nirenberg and Khorana.

DNA molecules sometimes change in odd ways. In 1946, a German-American biologist, Max Delbrück

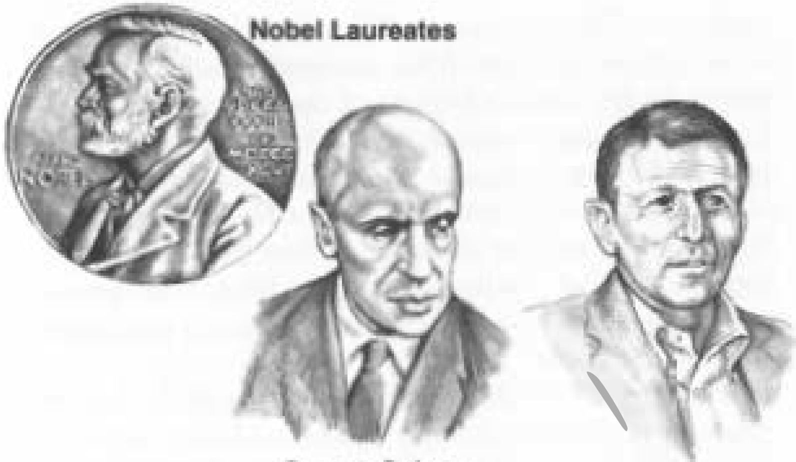
Nobel Laureates

Marshall Nirenberg
1968

Har Gobind Khorana
1968

Robert Holley
1968

Nobel Laureates



Severo Ochoa 1959

Paul Berg 1980

(1906–1981), and an Italian-American biologist, Salvador Edward Luria found that DNA molecules in viruses sometimes broke up for no reason that could be seen. A piece of DNA from one virus might then combine with a piece from another virus to form a different new virus. Hershey discovered this at about the same time. In 1969, all three shared the Nobel prize for medicine and physiology.

The question was could scientists do this on purpose? Could they break up DNA molecules and then recombine them in a different way?

Two American biologists, Daniel Nathans (1928–) and Hamilton Othanel Smith (1931–) discovered enzymes in 1970 and 1971 that could cause a break in a DNA molecule wherever a particular combination of nucleotides happened to be. In this way, the DNA could be broken into a few large pieces. If you knew the exact structure of the DNA molecules, you would

Nobel Laureates

James Watson 1962

Francis Crick 1962

Maurice Wilkins 1962

know exactly what pieces would be formed. Nathans and Smith were awarded the 1978 Nobel prize for physiology and medicine.

Soon afterward, an American biochemist, Paul Berg (1926–) found other enzymes that could recombine the pieces of the DNA molecule, but not necessarily in the original form. Two pieces from two different molecules could be recombined to form *recombinant DNA* that was different from any DNA found in nature. Berg was awarded a share of the 1980 Nobel Prize for chemistry for this.

By making use of recombinant DNA techniques, it is possible to breed bacteria, for instance, with new genes and new enzymes and new chemical abilities. Bacteria can be made to produce a hormone called *insulin* (IN-syoo-lin) that helps people with the serious illness, *diabetes* (DY-uh-BEE-teez). Animal insulin can be used, but bacteria can be designed to make human insulin, which is better.

Bacteria could be designed to produce other important materials, to convert pollution into harmless fluid, to perform certain useful chemical changes, and so on.

Of course, one must be careful. What if bacteria are produced that can give people brand-new and very serious diseases? There doesn't seem much chance of that, but biologists have to keep it in mind.

Would Miescher have dreamed that the strange compound he discovered a century and a quarter ago would turn out to be the key molecule in all living things from viruses to human beings? Would he have dreamed that human beings would learn to adjust and recombine those molecules and form the beginnings of a new science called genetic engineering?

Almost certainly not.