CHAPTER 14

THE CHEMICAL BASIS OF EVOLUTION

THE CHEMICAL BASIS OF EVOLUTION: GENES, CHROMOSOMES, AND MENDELIAN GENETICS

Science is an onion. It consists of questions, but each answer opens a new question, much as an onion consists of layer after layer of modified leaves. It is not unusual for a scientist to publish hundreds of papers over a lifetime, yet with each paper insist that he or she is working on the same problem—just delving deeper and deeper into the problem, peeling another layer off the onion. It is much the same issue as realizing the many layers of what seems to be a simple question that a child might ask. For instance, the child might ask, "Why are rabbits brown?" (Since science is concerned with mechanisms, rather than primary causes, "Why" is not an appropriate opening—see page 14—but, since children often use the expression, we will continue along this line.) One answer might be, "Because God made it brown," which is likely to lead to the follow-up question, "Why did God make it brown?" One answer might be, "To allow it to hide from its enemies." At this point the query could go in several directions—talking about natural selection, documenting that color does make a difference, considering the relationship of the color of the rabbit to the color of its environment and the presence of predators, discussing the genetics of pigmentation, or the biochemistry of the synthesis of pigments. If one followed the latter argument, the synthesis of pigments, this could lead to the question of why some molecules are transparent and others have colors, which could lead to an exploration of how atoms are held together into molecules, and how the interaction of atomic structure with light leads in some instances to the light's going through the atom or molecule and in other instances to the light's being absorbed or reflected. This is why the great thinkers of classical Greece, China, India, the Enlightenment, or other cultures did not resolve for all time the questions being asked.

This style is beautifully illustrated by the pursuit of the generalized question, "What is the basis of inheritance?" This pursuit led to the identification of DNA as the genetic material, and subsequently an understanding of how DNA carried information and how this information was transformed into the building blocks of all organisms. This is the story of the rise of molecular biology, surely one of the great episodes in the history of science. It is as abstruse and rarified as any level of knowledge today, but there is no reason why a student cannot understand how

it came about. The story illustrates spectacularly well how scientists ask questions and pursue them. For the most part, asking the question and getting an answer was a matter of games and tricks. In the vernacular, "Molecular biology ain't rocket science". It mostly is a matter of cool tricks.

REALLY COOL TRICK #1: HOW TO TURN A NOT-SO-BAD BUG INTO A REALLY BAD ONE. (THE GENETIC MATERIAL IS DNA.)

Once it was conceded that sperm and egg united to form a new individual with the characteristics of its parents, one had to ask, what was in the sperm and the egg that carried the characteristics. The egg contains yolk and nutrients for the embryo. The sperm is much simpler, containing mostly DNA and proteins, but there are many other components as well. At the level of biochemical skill available in the 1930's, even sperm were too complex to use to analyze this question. One needed a simpler model. This model came from microbiologists worried about how diseases were transmitted. In their pursuit of this question, they learned that bacteria could transmit characteristics from one organism to another, and that even dead bacteria could pass on their characteristics to living bacteria. Thus whatever carried the characteristic had to be chemical, and was not a "vital force" or other unique characteristic of living organisms.

The first assumption that everyone made was that the chemical was a protein. This seemed highly logical. Proteins are very complex structures, consisting of a string of a mixture of twenty different building blocks called amino acids, whereas nucleic acids are much simpler, being a string of only four types of their building blocks, called bases. If you consider the amino acids and the bases to be letters in an alphabet, an alphabet with 20 letters can produce a lot more words than an alphabet with only four letters. With an alphabet of twenty letters, we have a language. (This last sentence uses 15 of the letters of the English alphabet.) With four letters, we don't have much: four, fuor, foru, frou, fruo, rouf, ruof, etc. Even allowing words to be different lengths and allowing letters to be used two or three times in the same word (foor, fuur) the language is very restricted. As the logic went, only proteins have the complexity to store all the information needed to build an organism. Unfortunately, however, the first reasonable chemistry began to give a different answer.

The critical experiments, known as the Avery-MacLeod-McCarty experiments (really cool trick #1), were as follows:

Avery and MacLeod, working at what is now Rockefeller University, were studying a type of pneumonia caused by bacteria in mice. They were working with pneumonia-causing bacteria and had isolated a variant (mutant) strain that did not kill the mice. Bacteria can be grown in Petri dishes, in which each single bacterium multiplies and forms a single spot or colony on the dish, much as you may see mold springing up from several isolated spots on a piece of bread. The virulent or deadly form formed smooth colonies that looked like little droplets, while the non-virulent or non-killing form formed colonies that had rough, uneven edges.

We now know that the bacteria that form the smooth colonies secrete a somewhat gelatinous material that both creates the smooth appearance and protects the bacteria from attack by the mouse's defenses, its immune system. The bacteria of the rough variant cannot make this material and are quickly destroyed by the mouse's immune system.

Avery and MacLeod were trying to understand the difference between the rough and smooth bacteria, and what made the difference in the virulence. In one series of experiments, they injected a mouse simultaneously with boiled (dead) smooth bacteria and live rough bacteria. They found, to their surprise, that the mice died. When they took samples from the dead mice and cultured them, they found that what grew in the culture was smooth, virulent bacteria. This was tremendously exciting, because it meant that something from the dead bacteria could convert the rough bacteria into smooth bacteria. Since the experimenters could grow the bacteria and inject them into new mice, which subsequently would die, the rough bacteria had been permanently converted, or transformed, into the dangerous kind. Avery and MacLeod confirmed that they could not grow smooth bacteria from the boiled culture or cause disease if the dead smooth bacteria were injected alone. What this meant was that some chemical in the smooth bacteria survived and transformed the rough bacteria into smooth. It was not simply a question of the chemical coating the rough bacteria and protecting that generation of bacteria, since more smooth bacteria could be grown in the culture and could infect more mice. The rough bacteria truly had been transformed.

Now it was acknowledged that a chemical existed that could carry genetic information and transform one variant of bacteria into another. It now became possible to try to purify this chemical and, using the criterion of transformation, to identify what it was. The scientists pursued this goal and came to a surprising conclusion: the transforming material was DNA, not protein.

Nobody really believed them. DNA was far too dull and uninformative a molecule to carry information (remember our four-letter language). Besides, chemical methods weren't really that good, and even the purest DNA was contaminated with a few percent protein. Obviously, the real genetic material had to be a sort of super-protein that remained during the attempt to separate DNA from protein. The results were not dismissed out of hand, because after all there was nothing wrong with the data or the way the experiment was done, but no-one was really satisfied. The logic was not yet there.

REALLY COOL TRICK #2: BACTERIAL MILKSHAKES

The question still remained alive, until finally a refined means of doing this assessment was developed. This we can describe as the bacterial milkshake, or really cool trick #2. The issue was, can one get really pure DNA and protein, so that it is possible either to identify the "super protein" or confirm that the genetic material really is DNA? This question was addressed in 1952 by Alfred Hershey and Martha Chase. They took advantage of an interesting bit of biology and used

a kitchen trick to get an answer. The interesting bit of biology can be summarized in the well known poem, “Big bugs have little bugs/On their backs to bite’m/Little bugs have lesser bugs/And so on *ad infinitum*.” What this translates to is that even bacteria have parasites. The bacterial parasites are called bacteriophage, or “bacteria eaters”. These are viruses that attack bacteria, eat everything inside the bacterium, and produce new bacteriophage, or phage, that will attack other bacteria. They are quite vicious: if one has a “lawn” of bacteria (a thin layer of bacteria covering an entire Petri dish so that the whole surface is grayish) a single phage and its progeny will kill all the bacteria within range, creating a clear spot or “plaque” in the lawn (Fig. 14.1).

What makes this arrangement so interesting is the manner in which one type of phage attacks one type of common bacteria. This type of phage, which looks like a mini lollipop, attaches to the bacterium, stick end first. It injects something into the bacterium, leaving the shell of the lollipop on the outside. Obviously, what goes into the inside is what is the source of the new phage—in other words, the genetic material. What stays outside plays no further role, being abandoned with the bacterial membrane and wall when the dying bacterium bursts, releasing the new phage into the medium. The whole life cycle takes about 20 minutes. The question then is, what goes inside? To the scientist, the question is how to determine what goes inside.

If that question changes to whether it is DNA or protein that goes inside, there is a way to answer the question. DNA contains a lot of phosphorus but no sulfur, while

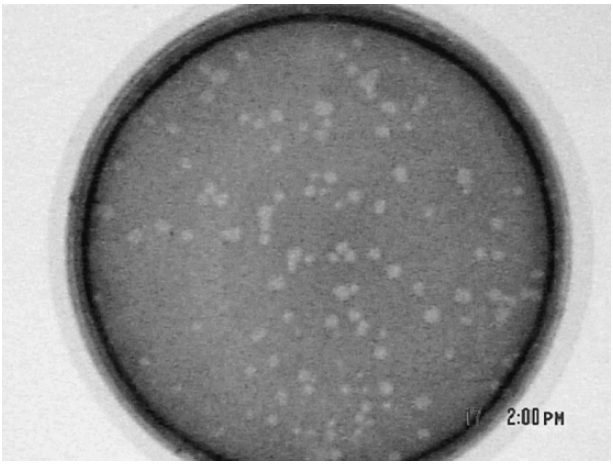

Figure 14.1. Bacterial plaques assay. In this experiment, viruses were scattered on a “lawn” of bacteria, otherwise described as an even coating of bacteria growing on medium in a Petri dish. The bacteria are stained and look dark in the picture. Where a virus has landed, it has infected the bacteria, grown, and reproduced, killing the host and infecting the bacteria next to it. Thus small circles of killing appear, as is marked by the clear areas, or plaques. Each plaque represents the descendents of one virus, or a clone of the virus. Credits: Photograph:–Jeffrey McLean, used with permission

proteins contain a lot of sulfur and not much phosphorus. By the time Hershey and Chase came along, the rise of the atomic era meant that reactors were producing, as byproducts, radioactive sulfur and radioactive phosphorus. Radioactive materials (radioisotopes) can be measured in extremely small amounts or, more importantly, trace contamination can be picked up at levels roughly 10,000 times less than can be detected chemically. So the contamination issue could be addressed, if one could separate what went in from what stayed out. This issue was handled with surprising simplicity. Simply put, the experiment was as follows: Phage were grown in the presence of both radioactive sulfur and radioactive phosphorus. These phage were used to infect bacteria. After a bit of time, but before the phage could kill the bacteria, the infected bacteria were thrown into an ordinary kitchen blender and blended. This knocks the phage off the bacteria. The mixture was then placed in a centrifuge, which spins the mixture at high speed, forcing all particulate matter to the bottom of the centrifuge tube. What comes down is the infected bacteria. What remains in the medium is what did not get inside and was knocked of the bacteria. Hershey and Chase, now having the outside of the bacteria separated from the inside, counted the radioactivity. The answer was unequivocal: The phosphorus (DNA) went in, while the sulfur (protein) stayed outside. The amount of protein that got in could be determined to be less than 0.1%. Thus it became almost impossible to maintain the argument for the "super protein", and the scientific world, reluctantly began to concede that the genetic material was DNA. The question now became, how did it work?

REALLY COOL TRICK #3: MOLECULAR BILLIARDS AND RUSSIAN DOLLS, OR DNA *MUST* BE THE GENETIC MATERIAL

There were many different directions from which one could attack this question, all of which had value, but the next cool trick was a molecular billiards game that made it intellectually necessary for DNA to be the genetic material. This was what is now known as development of the Watson-Crick model for DNA.

The expression "intellectually necessary for DNA to be the genetic material" is a bit of a tough nut to swallow, but it will make sense once we get to the end of the chapter. The first issue we have to address is how the structure of DNA was solved. Although the mechanism involves some of the most difficult aspects of biophysical chemistry, the principle is understandable if you take a little bit of time to think about it.

The chemistry of DNA was known. One can learn enough about chemistry to know that some reactions are possible and others are not. For instance, acids react with bases (vinegar reacts with baking soda) but acids do not generally react with each other. Iron reacts with oxygen to form rust, but gold does not. Using these kinds of arguments, chemists had deduced that DNA consisted of a long chain of sugar-phosphate molecules (deoxyribose phosphate) strung end-to-end:

(1) $\rightarrow$ sugar $\rightarrow$ phosphate $\rightarrow$ sugar $\rightarrow$ phosphate $\rightarrow$ sugar $\rightarrow$ phosphate $\rightarrow$

Simple sugars are made of a few carbon atoms. In the case of deoxyribose, there are five carbons per sugar. To one carbon in each of the sugars was attached a molecule of approximately similar size called a base. Since these did not form the backbone of the chain, they were considered to be side chains:

(2)
```
→ sugar → phosphate → sugar → phosphate → sugar → phosphate →
    |                   |                   |
  base                base                base
```

That was the chemistry. The question was, how did it actually fit together in space? To determine that, one had to get a good crystal of DNA, and then to apply some fairly straightforward tricks and thought to figuring out how it worked. Several laboratories tried to crystallize DNA, with Maurice Wilkins and Rosalind Franklin producing the best crystals.

Now came the question of what the crystal was. By a game of molecular billiards, it became possible to predict that the DNA chain was helical (Fig. 14.2). Linus Pauling had demonstrated a few years before that many proteins, which are chains of amino acids, took on a helical structure (the alpha helix, Fig. 15.1) and he had demonstrated how to recognize a molecular structure. You can understand the principle fairly easily. Throw two stones simultaneously into any suitable body of water, and watch the ripples, particularly where they meet:

As illustrated here, where the ripples meet, they will reinforce each other, producing a stronger ripple. Where the top of one ripple meets the bottom of another ripple, they will cancel each other out. Physicists describe this as the waves being in phase or out of phase. Looking at cross-sections of two waves, they would look like Fig. 14.3. When the waves are in phase, the result is a stronger wave (upper bold line). When the waves are opposite in phase (the trough of wave coincides with the peak of the other), the wave cancels out.

The point of this is that sound and light do the same thing. The "noise canceling" earphones that are sold for airplane travel cancel the sound of the engine by generating waves of sound out of phase with those produced by the airplane, and the shiny iridescence that one sees on puddles after a rain are produced by light reflecting off the top and bottom surface of a very thin layer of oil floating on the water. They can do that because the oil layer is just thin enough to cause the upper

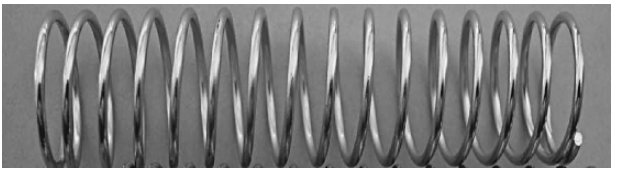

Figure 14.2. A helical structure. DNA is wrapped in a helical structure like this, with the sugar-phosphate chain forming the coils, and the coils held together by the interactions between the sugars and phosphates of one loop and the sugars and phosphates of the adjoining loops

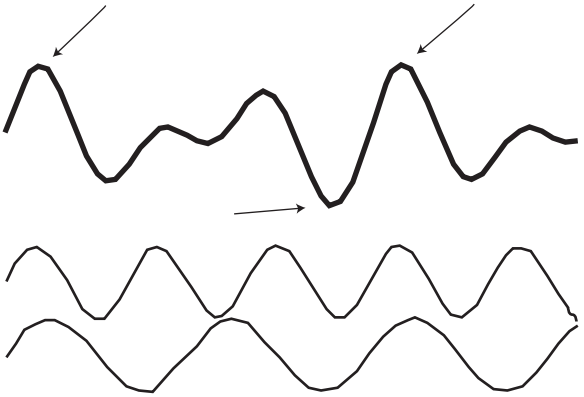

Figure 14.3. When waves meet (lower two lines), their heights sum (upper, heavier, line), so that when they are in phase at the peak or trough the combined wave is at maximum or minimum height (arrows). This phenomenon is noticed, for instance, in tuning a musical instrument. When the tone is close to that of the tuning fork, a beat is heard as the waves occasionally are in phase

and lower reflection to be out of phase. If light waves were the same length as the insides of molecules, we could see the equivalent of iridescence from molecules and be able to measure the molecule.

Unfortunately light waves are much bigger than that. However, the waves of X-rays, which in a sense are a much more intense form of light, are about the size that we expect molecules to be. If we then aim X-rays at a molecule, we might be able to say something about its structure. Specifically, if two successive waves can bounce off of two repeating units of a molecule, they will produce in-phase or out-of-phase reflections, depending on the distance between the repeating units.

This then is the molecular billiards game. The wave length of X-rays is known. When X-rays are shot at the crystal, the rays (billiard balls) bounce off successive repeating structures in the molecule. If X-ray film is placed at the proper point, where the X-rays (balls) hit in phase a spot will be produced, and its position will be a measurement of the distance between repeating units. The type of image that was acquired is illustrated in Fig. 14.4. From it, it was possible to conclude that the pattern was consistent with a helical structure. In other words, the X-rays were bouncing off successive loops of the spiral. The question then became, what did the spiral look like?

Helices can come in many forms, and it was important to understand what this one was. One clue was the density of the crystal. This can be explained as follows (Fig. 14.5): Suppose that you have a bunch of fairly loose springs, like the ones that contact the negative pole of batteries in portable electronic equipment:

You have a box full of them, which will represent your crystal. They can be all scattered loosely, in which case the box will weigh a certain amount, say one pound. It is also possible for the springs to be intertwined with each other. For

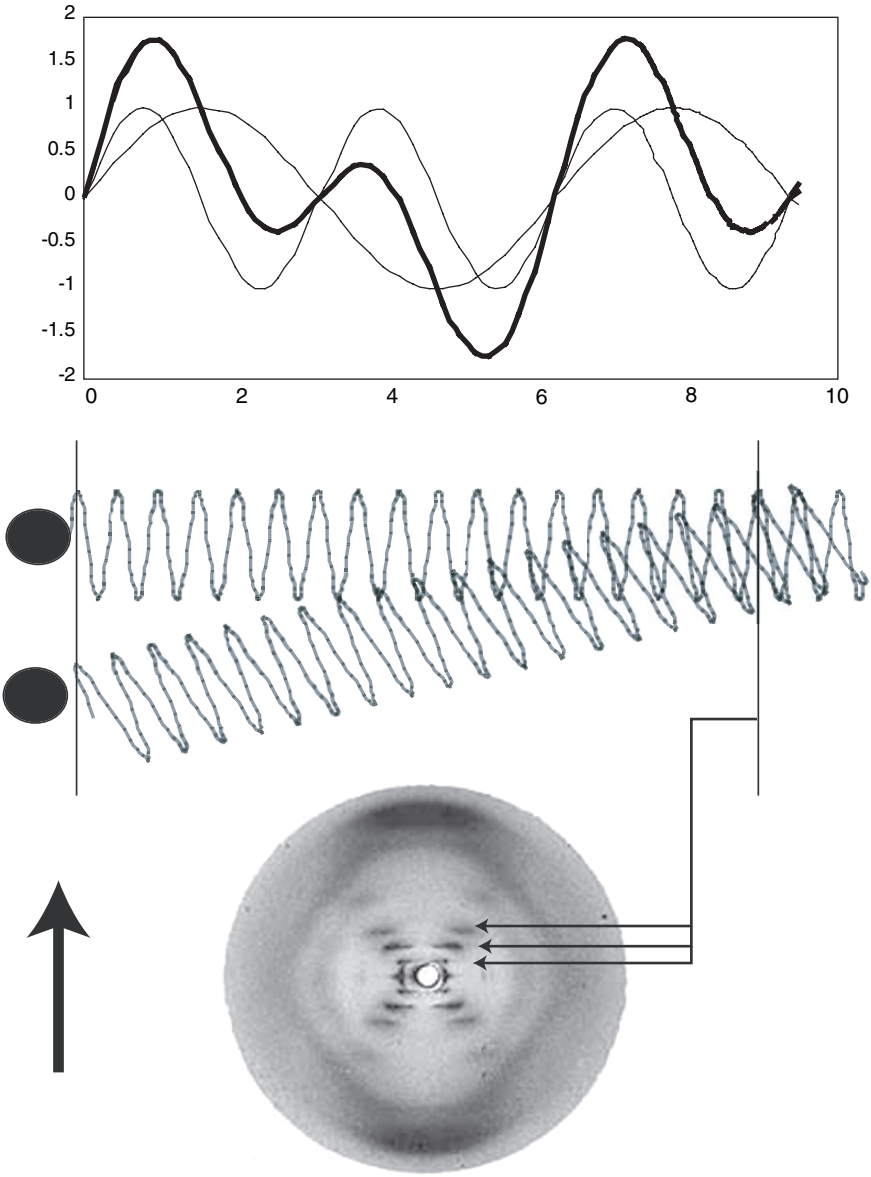

Figure 14.4. The effect of wave interaction. Upper panel: Waves reinforce when they are in phase (aligned with each other) and cancel when they are out of phase with each other (aligned opposite to each other. The heavy line is the resultant, or sum, of the two gray lines indicating waves of differing frequencies.

A crystal has a series of regularly-aligned atoms in it. For an X-ray hitting a crystal (arrow coming from bottom) the reflected waves go in all directions but start from slightly different positions. As they intersect, depending on whether they are in phase or out of phase, they will reinforce each other or cancel each other out. If a piece of X-ray film is set at an angle to catch the reflected rays, in-phase waves will

instance, you could have two springs or three springs in almost the same amount of space that you have one:

In the first case, the box would weigh two pounds, and in the second case, it would weigh three pounds. Since the box size is always the same, the density (weight per volume) doubles and triples. By weighing the DNA crystal and measuring its volume, it was possible to state that the crystal consisted of probably two and possibly three helices intertwined with each other, rather than one or four.

The final question, then, is how they fit together. Knowing the shapes of the sugar phosphates and bases, and the distances between the repeating units, Crick and Watson literally began to assemble models of how the different parts might fit together (Fig. 14.6). Among the various possible structures they found one that matched the numbers quite well. More importantly, it had constraints that led to the conclusion that we started with, that it was intellectually necessary that DNA be the genetic material. The constraints resulted from the measurements of the helix, which indicated that the DNA helix was actually two strands (a double helix). From their knowledge of the way in which the helix was constructed, they could identify both the pitch (distance from loop to loop) and the diameter of the helix. The constraint was imposed by the diameter. Again, based on the X-ray data, it looked like the bases (remember: the side chains) were on the inside of the helix, projecting into the center of the tube. However, the space in the center of the tube was not very generous, given the size of the bases. In fact, there were very few ways in which the bases could fit.

There were two limitations on the way in which the bases could fit. First, the bases are of two general types: a bulky form (purines) in which all the atoms form two rings attached to each other, and a smaller form (pyrimidines) in which all the atoms form only one ring. There was not enough room inside the ring for two purines to sit side-by-side. The only way it would work would be for two pyrimidines or one purine and one pyrimidine to sit side-by-side. Second, molecules can have local charges, vaguely like the north and south pole of a magnet, but here called positive and negative. They work like magnets, in that like charges repel each other and unlike charges attract. Two of the four possible bases are purines and two are pyrimidines. However, because of the way that the charges are distributed on the molecules and the way that they would fit inside the helix, not all combinations are possible. In fact, as Watson and Crick realized, there were only two possible combinations that would work: base A (adenine) across from base T (thymidine)

◄

reinforce each other and produce an exposed spot, and out-of-phase waves will cancel each other and not produce a spot. By knowing the length of the waves and applying suitable mathematics, it is possible to determine from the position of the spot the distance between repeating units, such as the loops of a helix. Lower panel: The X-ray crystalogram produced by Rosalind Franklin that ultimately was interpreted by James D. Watson and Francis Crick as representing a helical structure of DNA Credits: Franklin R, Gosling RG (1953) “Molecular Configuration in Sodium Thymonucleate”. Nature 171: 740–741

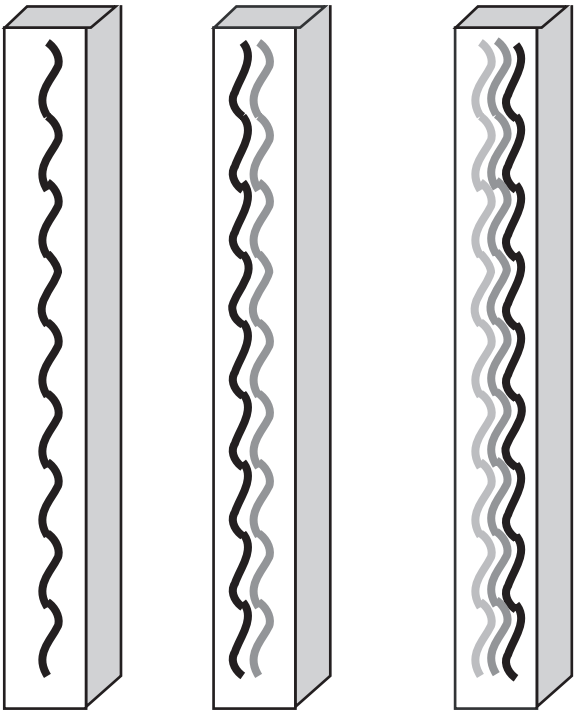

Figure 14.5. Density of crystals. Because helices can nestle in among each other, two or three helices can occupy a volume very similar to the volume occupied by only one helix. However, if one weighs each of these boxes, of course one gets different weights. Thus their densities, or weight divided by volume, are very different. If one can get a crystal of DNA large enough, one can measure its density. Another trick would be to suspend the crystal in liquids in which it will not dissolve, but of different densities. If it is less dense than the liquid, it will float, and if it is more dense, it will sink. This is the same type of analysis that Archimedes used to determine the amount of gold in the king's crown. (He jumped out of his bath and went running to tell the king, shouting "I found it!" [Eureka! In Greek].)

and base G (guanidine) across from base C (cytidine, Fig. 14.7). This would finally explain a curiosity known as Chargaff's rule, which stated that, no matter what the composition of the DNA, the amount of A always equaled the amount of T and G = C. But explaining this riddle was not the important issue. The pairing of A with T and G with C explained how DNA could be the genetic material and made it intellectually necessary for it to be so. In brief, if you pulled the two strands apart and rebuilt, for each strand, a new second strand, then the new second strand would necessarily be a duplicate of the strand that had been pulled off. If strand 1 had an A, then strand 2 had a T, and the new strand 2 (2a) would also have to have a T, while the new strand 1 (1a) built on the old strand 2 would have to have an A. In other words, each strand could create a new strand like the one that it had lost.

This resolved the Russian doll problem. You may know the Russian dolls, or matrioshkas, that come apart, revealing a smaller doll inside; the smaller doll also

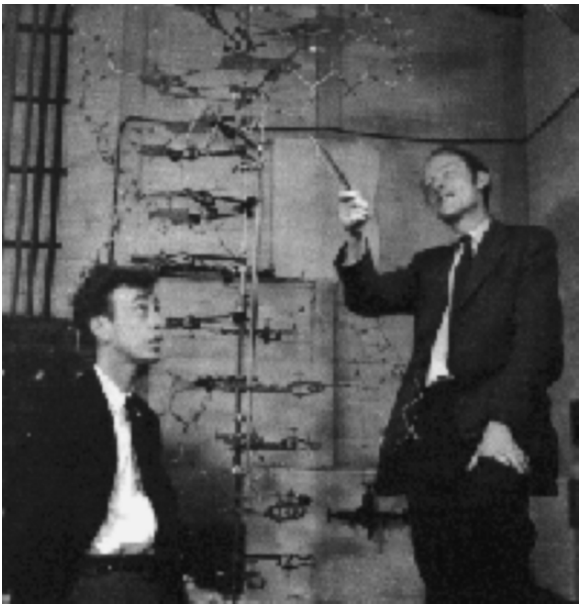

Figure 14.6. James Watson (left) and Francis Crick examine the model of DNA that they built to elucidate its structure (1953). Credits: Watson_and_Crick: library.thinkquest.org/C004535/nucleic_acids.html

comes apart, revealing a still smaller doll. In high quality dolls, there may be ten or so different dolls, one inside the other (Fig. 14.8).

In biology, the Russian doll problem consists of the following (supposing that the genetic material is protein, which forms the bulk of our bodies): If protein is the genetic material that carries the information for making (coding for) protein, what codes for the genetic material? In other words, what makes the protein that makes the protein that makes the protein? The Watson and Crick model demonstrated that, according to the structure of the double helix, each strand would serve as a template (mold) for a new strand WITHOUT HAVING TO HAVE FURTHER INFORMATION AVAILABLE. In other words, if the strands could be separated and a new strand could be assembled on each old strand, the molecule could replicate itself. This was the truly important element of the Watson-Crick model of DNA. They had identified a molecule that, by its structure, could be copied without having to have a code for the code for the code for the code for the code… Not only was it now possible for DNA to be the genetic material, because DNA gave an escape from the Russian doll problem, since no other molecule had this property, it was even necessary that DNA be considered to be the genetic material.

The question then turned to how it was possible for DNA to carry the information to produce a human or any other organism. Before we explore that question, however, you may want to note how many sciences ultimately

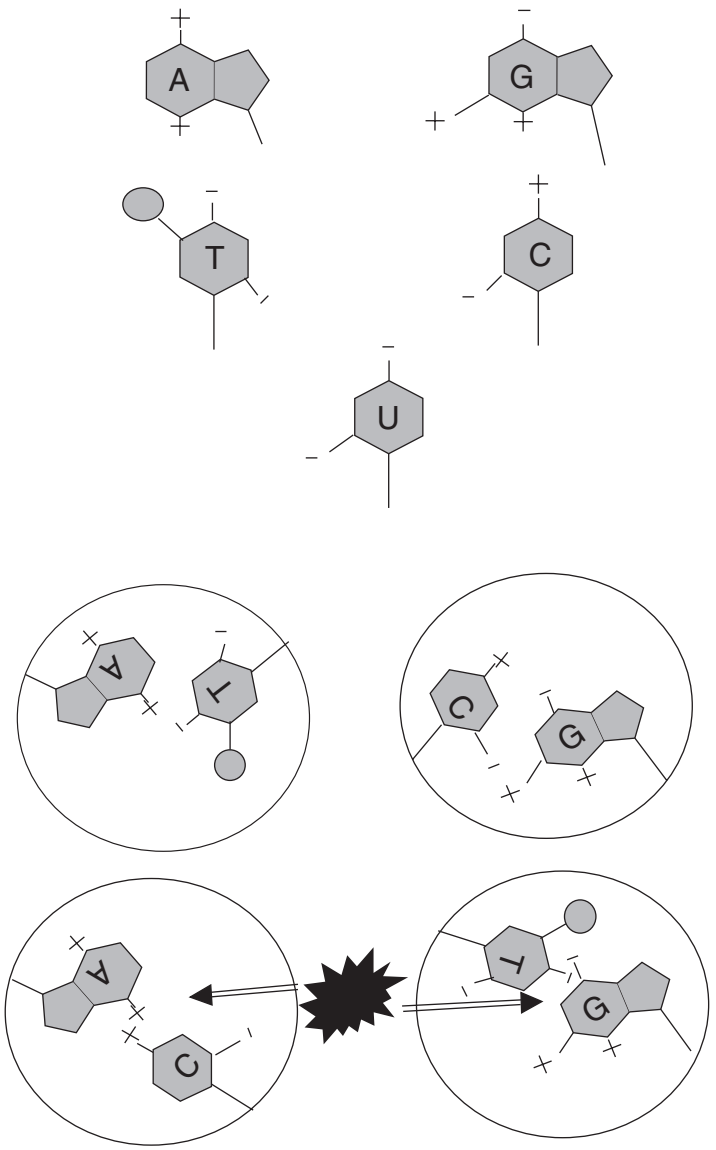

Figure 14.7. Purines and pyrimidines. Upper panel: The purines, A and G, are two-ring structures while the pyrimidines, T and C are smaller one-ring structures. U is used in RNA, while T is used in DNA. Thus the bases of DNA are A, T, G, and C, while in RNA they are A, U, G, and C.
Lower panel: Because of the sizes of the purines, two purines cannot fit across from each other inside the helix. Likewise, since charges on molecules act the same way that magnets do, in that like charges repel each other and unlike charges attract each other, C and A repel each other, as do G and T. T and C, although small enough to fit in the helix, likewise repel each other (not illustrated). This leaves only two possible combinations: A-T and G-C. All of these conclusions derive from the calculations of the size of the helix

Figure 14.8. Russian matrioshka dolls. Some of them have eight or ten dolls stacked inside each other. The question was, if protein carried the information to make protein, what carried the instruction to make the protein to make the protein that carried the information?

came to bear on this one question. Chemists had established reaction mechanisms and means of calculating and inferring the shapes of atoms and molecules; physicists had understood the wave properties of light and X-rays and how to interpret them; and biophysicists had learned to interpret complex patterns to reveal the structures of molecules. This is typical of any science, that each phase depends enormously on the work of predecessors, even in far-removed fields, and truly convincing arguments are based on the accumulation of data and understanding from many different fields. This is particularly true for the theory of evolution: As is noted on page 94 ff, the consistency of data from many fields is one of the strongest arguments possible for the argument that evolution has occurred.

Another point that you may care to notice is the following: we now can assemble a model of the molecule, and even see it in an electron microscope (another tool that contributes to our understanding). One of the most satisfying experiences in all of science is to see that an intellectual prediction turns out to be true (Fig. 14.9)

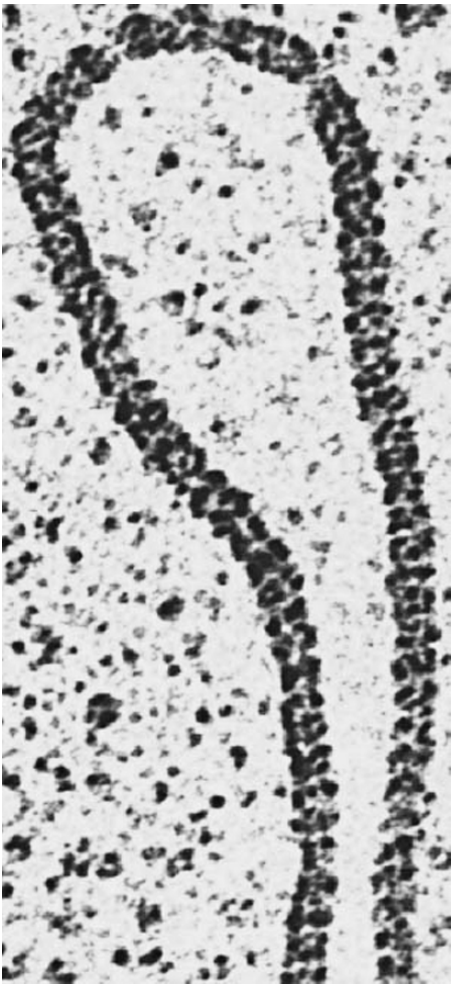

Figure 14.9. This loop of DNA has protein bound to it, as it is found naturally inside of cells, but the helical structure of the DNA strand is nevertheless clearly visible. Credits: Nucleic acids www.biochem.wisc.edu/inman/empics/Protein.jpg

So the question now becomes, "If DNA is the genetic material, how can it possibly carry information?" In other words, how can you possibly get something interesting from a four-color piece of string? Because that's what DNA was, a very long string with four variations. Very boring. Well, it is possible to get something more meaningful out of four variations, if you take the variations in groups. For instance, the Morse code consists of only two variations, dots and dashes, but by assigning values to sequences of one to four characters ("S" = dot, dot, dot; "O"=dash, dash, dash, etc.) one can create an entire alphabet. Another example

would be the rhyme heard in the US South, to help people distinguish between the similar-looking deadly coral snake and the harmless milk snake, by the sequence of three colors:

> "Red, black, yellow: Dangerous fellow. Red, white, black, that's all right, Jack"

DNA is a string of sugars with bases attached, and proteins are strings of amino acids. So it was logical to assume that the DNA string must somehow represent the protein string. It was already known that the genetic material must be arranged in linear order on the chromosome. This information was determined by very simple logic.

- Since the number of chromosomes is limited, there must be 1000 or more individual genes per chromosome.

- If different genes are on separate chromosomes, they will separate randomly, according to Mendelian genetics (pages 134 and 205).

- If different genes are on the same chromosomes, they should not separate at all, unless the chromosomes can break and rearrange (which they do).

- If the chromosomes can break and rearrange at random locations, then the closer two genes are to each other, the less frequently they should separate, in the same sense that, in a 1000-link chain, the chance of separating link 671 from link 672 in a random break is 1/1000 or 0.1%, while the chance of separating link 1 from link 1000 is 100%.

- One can determine the linear order of genes on a chromosome in this manner. In chromosomes that are big enough to analyze, such as those of the fruit fly *Drosophila*, the order is the same as the genetics indicates.

Francis Crick showed mathematically that if this were so, then it would take a sequence of three bases in a row to represent one amino acid. The math is very simple. If one base equals one amino acid, then there can be only four types of amino acids, but in fact there are twenty. If two bases in a row represent one amino acid, then there are sixteen possible pairs of the four bases—close, but no cigar. (Table 14.1) If three bases in a row represent one amino acid, then there are sixty-four possible combinations.

Thus three bases was the minimum number possible for such a coding to work. He even did an experiment to prove it. His hypothesis was that the linear string of DNA coded for the linear string of amino acids, with three bases in the DNA representing one amino acid. He also hypothesized that the code was read only by identifying the first base and stepping by three:

Table 14.1. Possible combinations from different numbers of bases

1 base	2 bases	3 bases			
A	AA	AAA	AAT	AAG	AAC
T	AT	ATA	ATT	ATG	ATC
G	AG	AGA	AGT	AGG	AGC
C	AC	ACA	ACT	ACG	ACC
	TA	TAA	TAT	TAG	TAC
	TT	TTA	TTT	TTG	TTC
	TG	TGA	TGT	TGG	TGC
	TC	TCA	TCT	TCG	TCC
	GA	GAA	GAT	GAG	GAC
	GT	GTA	GTT	GTG	GTC
	GG	GGA	GGT	GGG	GGC
	GC	GCA	GCT	GCG	GCC
	CA	CAA	CAT	CAG	CAC
	CT	CTA	CTT	CTG	CTC
	CG	CGA	CGT	CGG	CGC
	CC	CCA	CCT	CCG	CCC
4	16	64			

THEBADBOYFEDTHEFATCATANDDOGTHEBIGREDBUG
(THE BAD BOY FED THE FAT CAT AND DOG THE BIG RED BUG).

Therefore he proposed that getting out of sequence would be a disaster. It was known that certain chemicals could damage DNA (cause a mutation) by getting tangled in the helix and causing the DNA to add an extra base when it replicates, while radiation and other chemicals could damage a base and cause it to be lost. Therefore he proposed the following experiment: DNA was proposed to code for enzymes, proteins that can carry out reactions such as digesting food. If in bacteria he could cause a mutation by adding a base, the resulting enzyme would be a mess and would not work:

THEBAD**D**BOYFEDTHEFATCATANDDOGTHEBIGREDBUG
THE BAD DBO YFE DTH EFA TCA TAN DDO GTH EBI GRE DBU G..

Likewise, if he could cause a mutation by removing a base, the enzyme would not work:

THEBADBO.FEDTHEFATCATANDDOGTHEBIGREDBUG
THE BAD BOA NDT HEF ATC ATA NDD OGT HEB IGR EDB UG

However, if he could combine the two mutations, he might get an enzyme that would have one area that was a problem, but mostly it would be normal, and it might work:

THEBAD**D**BO.FEDTHEFATCATANDDOGTHEBIGREDBUG
THE BAD **DB**O FED THE FAT CAT AND DOG THE BIG RED BUG

He then did the experiment, and he got an enzyme that wasn't as good as the original, but did work. Thus the evidence supported the argument that the code was a sequence of three bases representing one amino acid, and the question turned to what the code was. To understand how that was done, we need to know a bit more about how it is possible to get the mutations that one wants to use to be able to examine a phenomenon. In other words, how could Crick get bacteria carrying precisely the two mutations that he would need to answer his question? The story of how this was done involves some of the coolest tricks that I know, which is the story of the origin of molecular biology.

REALLY COOL TRICK #4: PLAYING IN THE KITCHEN IS WONDERFUL FOR BABIES AND FUTURE NOBEL LAUREATES

Using mutants to study mechanisms was obviously a good idea—we can find out if this bulb is the blinking bulb in the Christmas string of lights by replacing it with a different bulb—but hoping to find the right mutation was not the way to go. Herman Muller improved the situation by showing that X-rays could cause mutations, and then producing a lot of them in fruit flies, but even fruit flies take two weeks to grow and require a lot of care. Bacteria grow very rapidly and cheaply (anyone who has let a bottle of milk spoil knows that you can get millions of bacteria in a quart of milk). They divide every twenty minutes. One bacterium will become one million in 20 generations, or less than 7 hours. It could theoretically become almost 5 trillion trillion billion ($4.7 * 10^{21}$ or 4.7 with 21 zeros) in a day. It won't of course. It would run out of food. Bacteria also have another advantage: they are haploid, meaning they have only one copy of each gene. Thus any mutation would be immediately obvious, as opposed to the situation for most diploid organisms, in which one characteristic may be hidden by another and be identifiable only by reproduction. For instance, if one of your parents has black hair and the other red hair, you could have black hair but carry the red hair characteristic hidden by the black, and no one would know that you had it unless one of your children or grandchildren was red-haired. However, there were two limitations to using bacteria to study genetics. First, the haploid style may be an advantage but can also be a disadvantage. Since bacteria don't have characteristics such as eye color or wing shape, most mutations that people can identify are the loss of the ability to use something, for instance the milk sugar lactose, as food. If one is looking for new mutations, the clearest evidence that one has found the mutation is that the bacterium has died, in which case of course the mutation has been lost: quite an embarrassment. Worse, in the early 1950's it was thought that bacteria did not recombine sexually. They were considered simply to keep dividing, replicating the same chromosomes over and over again. Thus getting mutations to study how things worked was an exercise

in frustration. If you wanted to ask, for instance, whether a bacterium's ability to resist penicillin was related to its ability to resist streptomycin, you could get a mutant that resisted penicillin and you could get a mutant that resisted streptomycin, but you would have no way to get both mutations in the same organism. Edward Tatum posed this problem to a young graduate student at Yale, Joshua Lederberg. Lederberg, literally playing around with kitchen equipment, figured out how not to lose a new mutation. By doing so, he demonstrated that bacteria could recombine successfully, and launched the era of molecular biology.

What Lederberg did was maddeningly simple, in the sense that brilliant experiments usually lead to a "Why didn't I think of that?" response. He made a rubber stamp. What he really did, as the story is told, is that he borrowed a piece of velvet from his wife. When you see a droplet of bacteria growing on a Petri dish or in a jar of jelly, what you are seeing is a colony of clones. One bacterium has landed there, found food, and kept dividing until there are hundreds of thousands or even millions of bacteria, each genetically identical to its parent, siblings, and progeny. If you touch a piece of velvet to the colony, the velvet will pick up some

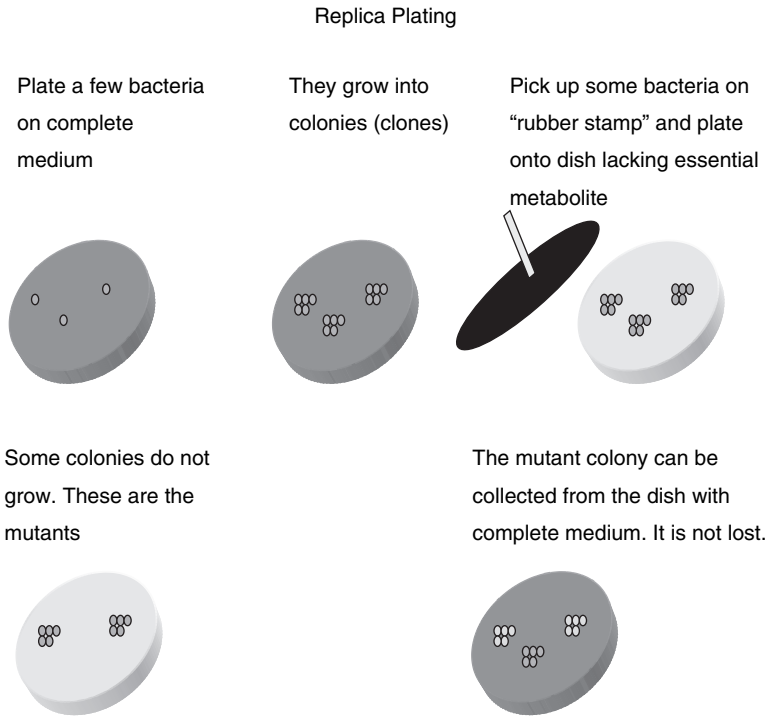

Figure 14.10. Replica plating. Since each colony represents the descendents of one bacterium, this technique provides a means of identifying mutants that cannot survive under certain conditions while not losing the mutant because it died. It was the key to the origin of molecular biology

of the bacteria. If you now touch that velvet to another Petri dish, you will leave some of the bacteria on the second dish, much as a rubber stamp leaves ink in the appropriate places on a piece of paper (Fig. 14.10). The genius of this experiment is that you can test for defects, such as the inability to make the amino acid arginine or tryptophan, by raising the bacteria on media lacking these ingredients, but you have not lost the original colony, which is still growing on the original Petri dish containing all possible nutrients. By using this trick, Lederberg was able to identify and collect many types of mutants. Others had suspected that such mutants existed, but had always lost them. The purpose of collecting the mutants was that Lederberg could now ask the basic question, could bacteria recombine sexually? It was the same question as asking if a lion and a leopard could mate and produce young, or if a peach and plum could be crossed to produce a nectarine.

Again, the basic experiment was very simple.

Lederberg had one mutant that could not produce the amino acid arginine. Let's call it arg–. Thus it could not grow in media lacking arginine. He knew that it could back-mutate only very rarely into a form that could produce arginine (arg+). Approximately one in 1,000,000 bacteria could do that. In other words, if he diluted the bacteria in a medium so that there were 10,000,000 bacteria per ml and spread that milliliter of bacterial suspension onto a Petri dish containing medium that lacked arginine, approximately 10 colonies would grow. Likewise, he had another mutant that could not produce another amino acid, tryptophan (trp–), and it could back-mutate at the same rate. He then mixed bacteria that could produce arginine but not tryptophan (arg+,trp–) with bacteria that could produce tryptophan but not arginine (arg–, trp+) and plated them onto a dish that contained neither tryptophan nor arginine. The only bacteria that could survive on this dish would have to be able to produce both arginine and tryptophan (arg+, trp+). This could arise in one of four ways: The arginine-requiring organism could back-mutate; the tryptophan-requiring organism could back-mutate; either one could produce a chemical that the other could use (this was ruled out by other experiments) or they could share genes, such that the arginine-deficient organism could get a good arginine gene from the tryptophan-deficient organism, and vice-versa. How could he tell?

The numbers gave it away. As we noted above, when arg– organisms were plated onto the arginine-lacking plate, only about one in one million could grow. The same result occurred if trp– organisms were plated onto the tryptophan-lacking plate. However, when he mixed the two types and plated them onto a dish lacking both arginine and tryptophan, one thousand colonies grew. In other words, by simply mixing the bacteria, he got a 100-fold increase in conversion. He grew these bacteria to show that this truly was an inherited difference, and otherwise eliminated the hypothesis that this might be chemical replacement of the missing nutrients. By eliminating all other hypotheses or interpretations, he was forced to the conclusion that mixing the two types of bacteria allowed them to exchange genetic material. In other words, bacteria could recombine sexually.

This was not simply a quirk or a silly story to tell at a party. It opened the possibility of moving genes around in bacteria to finally learn what genes were and

how they worked, which made it possible to do all the molecular biology that we do today. Furthermore, this recombination is the primary means by which bacteria develop resistance to antibiotics and multiple resistance (to many antibiotics). It also is a major means of viral mutation, and is a significant component in generating cancer cells.

Most of the mutations that Lederberg used were inability to make or digest products that the bacterium needed. In other words, they were failures of enzymes needed to synthesize the product or break it down into usable form. Enzymes are made of proteins, and we return to the question of how DNA carries the information. Once it was possible to produce bacteria with many types of mutations, it also became possible to ask how genes were constructed and how DNA carried information to make proteins. There are many stories about this search—most of which involve really cool tricks to get these molecules to reveal their secrets—but we cannot tell them all, and we do not have to maintain a strict historical sequence. Let us start with the question of learning how to identify the sequence of bases in DNA and learning how to read that sequence.

The first problem that we have to deal with is that there is a LOT of DNA. We have enough DNA to make 1,500,000 genes, though we actually have only 20–25,000 genes (the other 98.4% of the DNA being apparently useless used for instructions on when to be active or other, unknown functions) genes, and frogs have even more (Fig. 14.11). We are still at the level of trying to find out how three bases code for a single amino acid. Where do we start? The tricks that we pick up here are the same ones that will eventually be used for forensic analysis, for tracing the evolution of humans, for determining whether or not Neanderthals are related to us, and for genetic engineering, whether for crop production, repair of disease, or more dubious enterprises. We start by cutting the DNA up into manageable sizes, under controlled circumstances so that we know exactly where we are cutting it. This works because nature does it for us.

REALLY COOL TRICK #5: VIRUSES KNOW HOW TO CUT UP DNA

As we noted above, viruses are sometimes the diseases of bacteria—they infect and kill the bacteria. Sometimes, however, they do not kill bacteria but simply go along for the ride, like a parasite that does not really harm its host. The easiest way that they can do this is to hitch a ride on the bacterial chromosome. The bacterial chromosome, interestingly enough, is a circle; it has no end. What the virus does is open the circle, stick its DNA into it, and then close the circle back up, sort of the way that a magician makes two circles join together and come apart (Fig. 14.12). The bacterium then goes on with its life, dividing on schedule, but also replicating the viral DNA as it replicates its own.

For this to work, the virus has to cut the DNA in such a way that it will not disrupt an important function for the bacterium. It does this by identifying only very specific sites on the bacterial DNA where it will cut. This is accomplished because it has enzymes known as restriction endonucleases. The term “endonuclease” means that the enzyme cuts the DNA in the middle, rather than chewing in from the

end, and the qualifier "restriction" means that at a specific sequence of DNA, for instance a specific sequence of four to six bases, for instance GAATTC. Even more interesting, look again at the sequence. It is what we call a palindrome. Palindromes are sentences that read the same backwards and forwards: MADAM I'M ADAM or ABLE WAS I ERE I SAW ELBA. In this case, the palindrome is the opposite strand, which reads backwards exactly like the first strand:

GAATTC →

← CTTAAG

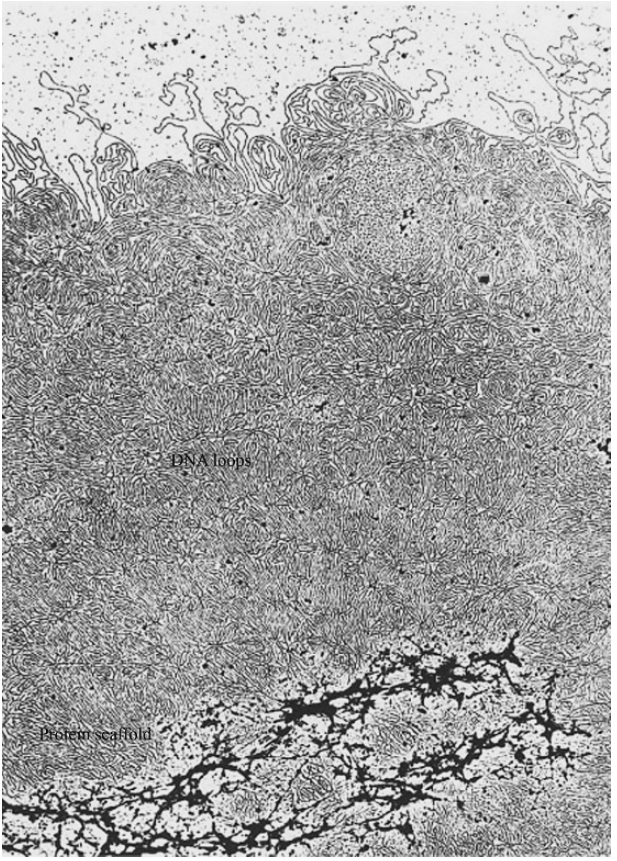

Figure 14.11. There is a lot of DNA in a cell. In this preparation, a mitotic chromosome was spread on the surface of water to allow it to expand. All of the fine strands are DNA. This article was published in Cell, Vol 12, J.R. Paulson and U.K. Laemmli, The structure of histone-depleted metaphase chromosomes, Pages 817–828, Copyright Elsevier (1977). Credits: From J.R. Paulson and U.K. Laemmli, 1977. Cell 12: 817

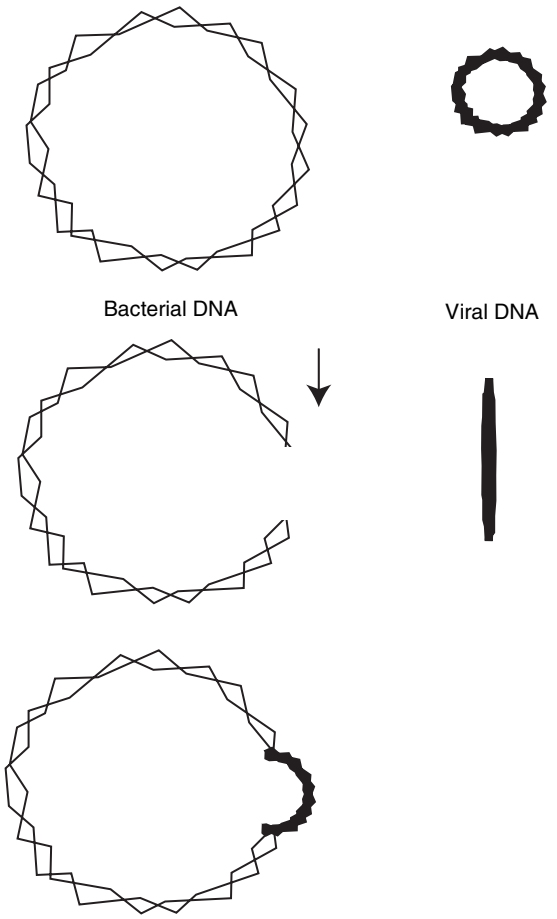

Figure 14.12. Viral insertion into DNA. The DNA of both the bacterium and the virus (phage) are circular. The virus cuts the DNA of the bacterium and simultaneously opens its own DNA into a straight piece of DNA. It then attaches the ends of its DNA to the ends of the bacterial DNA and splices the circle back together. As the bacterium reproduces, it makes a copy of the viral DNA as well as a copy of its own

This is fine, because the strands face in different directions. (In the sugar-phosphate-sugar-phosphate backbone, the phosphate is attached differently to the two sugars. It's rather as if you had two battery holders, each of which took a string of batteries, all facing the same way, but the battery holders were set up so that in one, all the positive poles faced left and in the other, they faced right.) If, for instance, the restriction enzyme cut between the G and the A in this sequence, as the enzyme EcoR1 (for E. coli Restriction Enzyme 1) does, then it will cut both strands, leaving a little bit dangling over. The dangle will become very important in a bit.

Consider what this means: The probability of finding an A next to a G is 1 in 4; the probability of finding AA next to G is ¼ × ¼, or 1/16. The probability of finding the entire sequence is ¼ × ¼ × ¼ × ¼ × ¼, or just about 1/1000. Fruit flies have about 122 million bases and 14,000 genes, while humans have about 3 billion bases and 20 to 25,000 genes. This means that this one enzyme might cut up human DNA into 3 million pieces. If we can use it, it would be like taking a very unwieldy book with no punctuation and cutting it into pieces by cutting every time we found the ending "-ation". (It would be even more meaningful if we cut it every time we encountered the word "chapter".) If we can separate the DNAs from different chromosomes or by other characteristics—this can be done—we can get a manageable number of fragments to analyze. The restriction endonuclease is the first of several tricks in this bag. It is now used commonly in forensic medicine. This is how it is used:

There are many regions of human DNA that are very variable, so much so that they are nearly unique for every person. If we can analyze that region, which we can identify by another trick, we can distinguish one human from another. Restriction endonucleases come into play because the piece of this variable region will be the same size only if the two pieces we are comparing are identical. Look at the following sentences, from which we will cut a piece by cutting only after the string of characters "and the":

The buffalo and the *prairie dog are characteristic of the plains. The cockroach and the* pigeon are characteristic of the city. (69 characters)

The buffalo and the *prairie dogs are characteristic of the plains. The cockroach and the* pigeon are characteristic of the city. (70 characters)

The buffalo and the *prairie dog are characteristic of the plains. The cockroach and pigeon are characteristic of the city. XXX* (≫103 characters)

The buffalo and the *prairie dog are of the plains. The cockroach and the* pigeon are characteristic of the city. (deletion: 57 characters)

The buffalo and prairie dog are characteristic of the plains. The cockroach and the *igeon are characteristic of the city.* Xxxxxxxx (?? characters)

Thus it is clear that, if we can separate pieces of DNA by size, we can identify them, analyze them, or at least distinguish which are identical and which are not. How to separate them by size is the next cool trick. It depends on the same principle that can be seen in mid-afternoon in major cities: middle school children get through the subway turnstiles much faster than big or even obese people. That's really cool trick #6.

REALLY COOL TRICK #6: GETTING DNA TO RACE

Everyone is familiar with gels such as gelatin ("Jello" ®) and you may have seen agar on a Petri dish. These gels are made of long strands of molecules (proteins or carbohydrates) that are tangled among each other, rather like a bowl of spaghetti. However, they also hold onto water very well, so that the structure of the gel is water suspended among the strands as if it is in a sponge. What is interesting about this is that the meshwork of strands leaves holes (water passages) about the size of

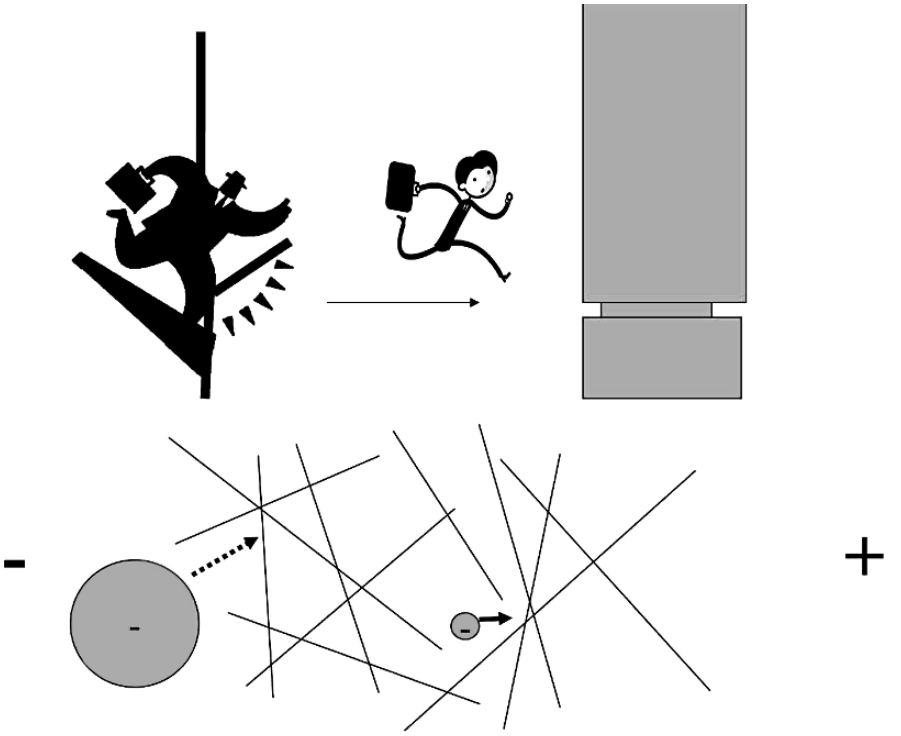

Figure 14.13. Electrophoresis. In the same way that, in racing for a subway car (gray) a small child can squiggle through the gate faster than a much larger individual (upper panel), small molecules can squiggle through gates faster than large molecules (lower panel). A gel is made such that it has holes equivalent to gates, and the holes are approximately the size of molecules. Since most proteins are negatively charged or can be made to be negatively charged, they can be attracted to the positive pole (anode) of an electrical field. However, since they have to cross the gel to get there, the smaller molecules move faster. Thus, proteins can be separated, and eventually identified, by size

molecules. By increasing or decreasing the amount of the material to make the gel, we can get gels with holes of different sizes. We can use this to make gates for the pieces of DNA.

Happily, DNA is an acid (deoxyribonucleic acid) and a characteristic of acids is that under the right conditions they are negatively charged, or negative ions. If you place ions between the positive and negative poles of a battery, negative ions will move toward the positive pole and positive ions toward the negative pole. So, if we put DNA between two electrical poles, but interpose our gel between them, the DNA molecules will move toward the positive pole, but the smaller ones will move through the gel faster, while the bigger ones get tangled in the mesh. This is electrophoresis (Fig. 14.13).

This would be fine if we could get all the DNA we wanted, but often it is hard to come by, from crime scene evidence, or a dinosaur bone, or from a newborn mouse

that we would like to identify but that we do not want to kill to get its DNA. Because DNA of interest was often in very short supply, Kary Mullis searched for a way to get more of it, and developed really cool trick #7, the polymerase chain reaction (PCR).

REALLY COOL TRICK #7: LOOKING FOR CRAZY BACTERIA

If you have ever tried to untangle two springs, a Slinky® toy, a hose, electrical cord, yarn, or a braid, you know that rotating one strand causes the other to rotate as well. It actually is quite a complex trick to unwind DNA, but in order for DNA to replicate, the strands have to separate so that a new strand can be built on each old strand. All organisms do this by using a complex set of enzymes collectively called DNA polymerase. The polymerase recognizes single-stranded DNA and builds a new strand on it. DNA will also unwind at high temperatures, so theoretically you could use that unwound DNA to make new DNA, but unfortunately the polymerase is cooked at that temperature. Proteins including enzymes are not stable at high temperatures. They get permanently deformed and they precipitate, as the white of an egg (mostly the protein albumen) does when you cook it. Thus we have a problem: We can get the DNA unwound, but we cannot use it at the temperature at which it is unwound.

There is, however, a solution. There are organisms (bacteria) that live in hot springs, such as those at Yellowstone Park. Some bacteria live in water hotter than 90° C (194° F)! If they live at that temperature, then it follows that they reproduce at that temperature, meaning that their DNA polymerase can survive and work at high temperature, the temperature at which most DNAs naturally unwind. Using their polymerase (Taq polymerase, from the bacterium *Thermophilus aquaticus*: Translated from the Greek and Latin roots, the name means "heat-loving creature in the water"), we might be able to synthesize new DNA from the strands of the unwound DNA. That is the "P" (polymerase) of "PCR". But it only doubles the amount of DNA. Doubling the amount of a vanishingly small amount of DNA does not help much. That's where the "CR" (chain reaction) comes in. If we run the cycle once, the 2 original strands become 4. If we run the cycle a second time, the 4 strands become 8. If we run it a third time, the 8 strands become 16. In other words, something that doubles each time increases at a geometric rate. If we run the cycle 10 times, we have increased our original DNA 1000-fold. If we run it 30 times, we have increased it one billion-fold. This is the normal procedure for PCR: by an automated procedure, a trace amount of DNA is run through approximately 30 cycles of polymerase reaction, thereby creating enough DNA to work with and study. Of course, everything depends on having a really clean bit of original DNA at the start and not getting fingerprints, dust, or bacteria into the preparation. You really don't want a one-billion fold amplification of that hamburger grease that was on your fingers.

Finally, to study the evolution of animals, one very important tool today is to compare sequences of DNA from different animals to see how closely they are related. For instance, we know that the gene for our hemoglobin is extremely similar

to that of chimpanzees, less similar to that of other mammals, but more similar to that of mammals than to that of birds, etc., and we can trace resemblances all the way to fish and beyond. Using the DNA, we can determine where whales came from (hippopotamus-like animals) and where vertebrates arose (from starfish-like animals). We can trace human migrations, and, making some assumptions about how fast mutations arise, use the number of mutations as a molecular clock (see Chapter 9 page 118). To do so, we need to be able to sequence the DNA. It turns out that this is really quite easy. We can do it by electrophoresis as we described above. Of course, we need a couple of tricks.

REALLY COOL TRICK #8: IF THE WATER MAIN IS WORKING AT 1ST ST. BUT NOT AT 3RD ST., THEN THE BLOCKAGE MUST BE NEAR 2ND ST.

DNA polymerase works by adding one base at a time to a growing chain bound to the intact chain. As we have discussed, the backbone of the chain is sugar-phosphate-sugar-phosphate-sugar-phosphate…The phosphate links to the sugars through oxygen on the sugar. (A simple sugar, like grape sugar, consists of 6 C, 12 H, and 6 O. Table sugar consists of 12 C, 24 H, and 12 O. To make a chain of sugars, the phosphate links to an O on one end of one sugar to an O on the other end of the next sugar.) If one of those oxygens is missing, the phosphate cannot link like a series of hook-and-eye links, with one hook missing, and the extension of the chain will stop. There is a synthetic form of base like this, called a dideoxy base (deoxyribose, the sugar of DNA, already lacks one oxygen; dideoxyribose also lacks the oxygen to which the phosphate would bind). A chain terminated by the addition of a dideoxyribose will be shorter than normal, and will run faster in electrophoresis. We could therefore recognize its existence, but how do we know what it is, and how do we read sequences?

There were several efforts to resolve this problem, but the one that has worked very well is this: It is possible to make a dideoxybase fluorescent and, better yet, make each of the four (A,T,G,C) fluorescent in a different color. What we now do is to prepare our DNA-synthesizing mixture with DNA polymerase and the DNA we wish to sequence. To this mixture we add a mixture of the fluorescent dideoxy bases, but not enough to stop the reaction totally. Let's see what happens. We will assume that the strand synthesis begins with a G. Some of the strands (there are actually millions of separate strands) will incorporate the fluorescent dideoxy G (ddG) and therefore end. Thus the shortest strand will have the fluorescent ddG plus the base that it was attached to. Other strands will incorporate a normal G and go to the next base—let's say it is an A. Some strands will incorporate the ddA and stop. Therefore the second shortest strand will be the unknown base-G-ddA and be three bases long. Other strands will incorporate a normal A and continue. After a while, we will accumulate a series of newly synthesized DNA, each type of strand being one base longer than the previous, and each ending with a fluorescent dd base. Let's electrophorese this, and put a sensor (light meter)

along the path. The light meter is capable of distinguishing the colors of the bases. It now records the strands as they pass by: G, A, …. In other words, the fluorescent bases are establishing the sequence for us! In reality, it is not possible to separate more than about 1000 bases at a time. To sequence entire genomes (all the genes) of animals, the DNA is broken into small pieces by restriction endonucleases, the sequences of the pieces are read, and the continuation of one piece to the next is identified by overlapping pieces, as you might fit together a torn-up newspaper by pairing partial letters from one piece with partial letters on the next piece.

The final really cool trick is described more because it is headline news than because of its relevance, but rearranging genes can be used to study evolution. It is the principle of genetic engineering, and we are aware that virus invasion of our chromosomes has changed our inheritance. The viruses use this trick, and it is the basis of all the stories of headlines. It relies on the palindromic sequences described above that some restriction enzymes use.

REALLY COOL TRICK #9: WHEN ALL THE PUZZLE PIECES ARE THE SAME

DNA can be damaged in many ways—by sunburn, X-rays, heat, and many chemicals. In order to survive, all organisms must have means of recognizing and repairing the damage where possible. One of the means of repairing DNA broken at the sugar-phosphate bond is an enzyme called DNA ligase. It can be isolated, purified, and used in the laboratory.

Now look at the palindromic sequence produced by EcoR1:

```
GAATTC →
← CTTAAG
```

The cut comes between the G and the A, leaving the two strands as

```
xxxxxG          AATTCxxxxxxxx
yyyyyCTTAA          Gyyyyyyyy
```

Note the AATT loose ends. Remember that any cut that the endonuclease makes will produce these loose ends. These loose ends, or overhangs, can still stick together, with the A's and T's still associating or binding, and the DNA ligase can repair that kind of break. But suppose that the strand on the right side came from a different piece of DNA?

```
xxxxxG.AATTCwwwwwww
yyyyyCTTAA.Gzzzzzzz
```

The DNA ligase would not be able to distinguish between the "good" DNA and the "fake" DNA; all it would do would be to repair that break, and the DNA would be a hybrid of the original piece on the left and the original piece on the right—an engineered piece of DNA. This is the heart of "genetic engineering". Different pieces of DNA can be attached to each other so that, for instance, a gene conferring resistance to frost can be inserted into a crop plant, allowing the plant to be grown in more northerly areas. Some of the better-known agricultural uses today include adding growth hormone genes to farm-raised fish to increase their growth rate and causing some crop plants to automatically produce insecticides normally produced by other plants or bacteria. Medical uses include production of usable quantities of hormones by cells grown in culture, production of highly specific and highly sensitive diagnostic reagents and production of specific proteins, sometimes deliberately altered, to fight specific diseases or cancers. The bulk of the most exciting advances in biomedical research today are based on the use of animals and plants with manipulated genes. There are as yet no cures based on "correcting" genes in individuals, because it is one thing to get a cell in culture to produce a specific protein, but it is much more complex to assure that one can place a specific cell in the body, at a specific location, so that it will produce the desired protein product only when it is needed and will distribute it only where it should go. There are also dangers inherent in altering organisms, mostly related to their potential to escape and compete with the native forms. Laboratory animals typically not only carry the desired altered DNA but are bred so that they cannot survive in the wild, but it is not guaranteed that all agricultural restrictions are so stringent.

However, the dangers are often greatly exaggerated. Essentially no food that you eat today resembles its wild form. All have been manipulated by selective breeding, deliberate induction of mutations, and cross-breeding, to improve the size, palatability, or appearance of the food. The original tomato was much smaller and berry-like, similar to its relative the nightshade. Potatoes, also members of the nightshade family, could be very toxic if they were allowed to become green or to sprout, but the toxicity has been bred out of them. Corn was similar to a tassel of grass. No apple or peach that you eat today resembles the original crabapple-like fruit but instead is a sterile hybrid, propagated by grafting onto other roots. All the variations of oranges have a similar history, as is indicated by names indicating human oranges: tangerine (from Tangiers); Clementine (after St. Clement); mandarin (from China). The large and almost always successful production of wheat in the western world depends on the judicious choice of insect-resistant strains when insects are a major problem, rust-resistant strains when this fungus spreads, and wheats specifically selected to emerge early in the spring or to grow late into the fall. Genetic engineering is a more efficient means of doing what we have always been doing, but does not represent a theoretically or morally new direction in human behavior.

REFERENCES

Campbell, Neil A. and Reece, Jane B.(2004) Biology (7th Edition), Benjamin Cummings, Boston, MA.

STUDY QUESTIONS

1. For any of the examples given above, describe the ELF logic on which it is based. Are there any flaws or limitations to this logic?
2. Make your own diagram of the several steps necessary to isolate and sequence a specific piece of DNA. Explain these to a classmate.
3. Make your own diagram of the several steps necessary to introduce a new piece of DNA into another piece of DNA. Explain these steps to a classmate.
4. Which of the "really cool tricks" do you find to be the most intriguing? Why?
5. Do you think that "really cool tricks" were used to build other sciences? Why or why not?
6. Describe a situation in which you or a friend or relative worked out a clever or ingenious means of solving a particular problem. Was this solution effectively different from the "cool tricks" that eventually led to Nobel Prizes? Why or why not?
7. Some scientists claim that, "For every difficult experiment, there is one organism that will be perfect to conduct the experiment." Does the story of molecular biology support or contradict this claim?