The Process Whereby Your Genes Make Your Proteins

You don't have to have a science degree to understand the means by which your genes influence your health. You do, however, need to know a little bit about the process by which your genes make your proteins. Once you understand the process whereby your genes make your proteins, you will be able to understand how the fact that different people have different sequences for a typical gene translates into the fact that different people have different levels of susceptibility to any given disease, or that different people will respond differently to a given drug.

First, A Word About Cells And Proteins

Your Body Is Like A City

Your body is a lot like a city. It contains approximately 100 trillion (100,000,000,000,000) cells that live among each other like city neighbors do. Each cell is a self-contained unit. If you peel a sunburn, or shave a (very thin!) shaving of skin off a callous in your hand, and look at it under a microscope, you can see the individual cells, lying side-by-side, with obvious boundaries (called cell membranes) that separate a cell from its neighbors. It looks similar to the scene you see when you look down at the ground from an airplane. You'd see the individual properties, with their boundaries, lying side-by-side, the way the cells do in your tissues.

One of the reasons why we humans are such complicated creatures is that we have approximately 210 different types of tissues, each of which can perform a specific function. Cells that perform different functions live among each other, just as different people who live in a neighborhood hold different jobs. For example, consider the tissues that make up your stomach. Your stomach has smooth muscles that contract so your stomach can mash up your food, as well as nerves that cause the smooth muscles to contract when you have food in your stomach. In addition, there are cells in the lining of the stomach that secrete hydrochloric acid, which helps digest proteins, other cells that secrete mucus, to protect the stomach's tissues from the hydrochloric acid, and still other cells that secrete other things you need for proper digestion. Along with that, you have some tough connective tissue. This tissue weaves among all the other tissues, holding them all in place together, and keeping the stomach from tearing up when it contracts and mashes up your food.

If any one of those tissues is not working properly, your stomach will not digest your food as you need it to. You will lose weight, feel nauseous, and maybe even have diarrhea. Your body needs all these neighboring cells to work together in a coordinated fashion in order for your organs to work properly, just like a city will not be able to maintain its roads and buildings, provide police and fire services, and import the food, clothing and other goods its people need for their daily lives if its people do not perform the right variety of functions, such as construction workers, police/firefighters, truck drivers and store owners.

The proteins that a cell has are like the different workers you see bustling around a city. Each protein has a specific function, and between them, your proteins do most of the work that is required to keep you healthy. Enzymes digest your food, metabolize your medicines and detoxify dangerous chemicals. Transport proteins enable you to take up nutrients from your food, or medications you have taken by mouth, and distribute these essential things to the tissues that need them. Antibodies and other immune system proteins fight against bacteria and viruses that would make you sick. DNA repair proteins prevent cancer by repairing the damage that smoking, sun exposure or barbecued meat do to your DNA. Just as a city needs all its different workers, with their unique functions, to work in concert with each other, you need all your proteins, with their unique functions, to work in concert with each other if you are to stay healthy.

The Level Of Activity In A Typical Protein Varies Between Individuals

If you understand that your proteins perform almost all the functions that are required to keep you healthy, you can see how, if two people have different levels of activity in a protein, they will have different levels of susceptibility to diseases, or respond differently to a prescription drug. For example, if two sunbathers have different levels of activity in one of the proteins that repairs damaged DNA, the sunbather with the lower level of activity will have a higher risk of getting skin cancer than the sunbather with the higher level of activity in the DNA repair proteins. Similarly, if two people have different levels of activity in a protein that breaks down a prescription drug, the person with the lower level of activity in that protein will build up a higher concentration of the drug in his/her blood than the person with the higher level of activity in that protein will. Because the risk of toxic side effects increases as the concentration of the drug in your body increases, the person with the lower level of activity in this protein will have a greater risk for having a toxic side effect from the drug than the person with the higher level of activity in this protein will. On the other hand, the person with the higher level of activity in this protein may fail to respond to a dose of the drug that works for the other person, because he/she clears the drug out of his/her body so quickly that he/she never builds up enough of a concentration of the drug for the drug to have its effects.

One of this book’s most important goals is to help you understand how the fact that different people have different levels of activity in the typical protein translates into the fact that different people have different levels of susceptibility to diseases, or different responses to drugs. In order to understand why different people have different levels of activity in the typical protein, you must first understand the process whereby a gene makes its protein. In addition, understanding the way in which your genes make your proteins will help you understand what the genetic tests we discuss later actually look at, and why these tests can provide useful information regarding the functioning of some of your proteins.

The Process By Which Your Genes Make Your Proteins

You don’t need to understand all of the details of the process whereby a gene makes a protein to understand how your genes affect your health. You do need to know a little bit about the process, however. Don't get bogged down in the details. Keep this overview in mind:

- The gene’s DNA provides the instructions to make the gene’s messenger RNA (mRNA)
- The gene’s mRNA provides the instructions to make the gene’s protein
- The cell makes the gene’s protein by chaining a set of amino acids together. The mRNA tells the cell which specific amino acids to chain together to make that gene’s protein.

The Gene’s DNA Provides The Instructions To Make The Gene’s Messenger RNA (mRNA)

Figure 1 is intended to help you picture the structure of the DNA molecule. You've no doubt heard the term "double helix" used to describe the structure of your DNA molecule. You may not have heard the term "nucleotide," however, but nucleotides are the building blocks of your DNA. DNA is made up of two strands of nucleotides. Because of the nucleotides' chemical structures, when you line up two strands of nucleotides, the complex adopts a helical structure, like a spiral staircase.

As panel A in Figure 1 illustrates, each nucleotide in your DNA is made up of three components: the sugar deoxyribose (shown as a pentagon), a phosphate group (shown as a circle with the letter “P” inside) and a base. There are four different types of nucleotides in DNA; each one contains a different base. Some nucleotides contain the base adenine (A), some contain the base cytosine (C), some contain the base guanine (G) and others contain the base thymine (T). Panel B in Figure 1 shows how the sugars and the phosphate groups of the nucleotides make up the backbone of the double helix, while the bases project inward from the backbones. Bonds between the bases hold the two DNA strands together. The two DNA strands line up in an **antiparallel** arrangement. This simply means that the two strands are parallel to each other, but oriented in opposite directions. This is similar to the way a two-lane road is laid out; the two lanes are parallel to each other, but the flow of traffic goes in opposite directions in the two lanes. Panel C illustrates the fact that A and T bases line up across from each other, while C and G bases line up across from each other.

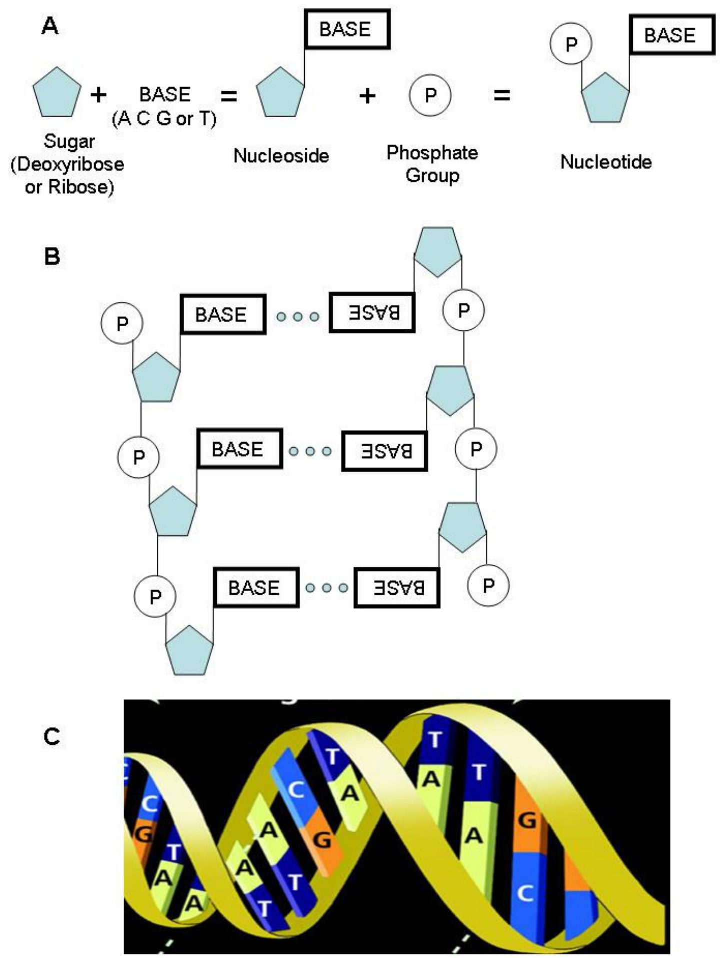

Figure 1. A) The three components of a nucleotide, B) the arrangement whereby nucleotides are joined together to make the DNA double helix, C) A bases bond with T bases, and C bases bond with G bases. Figure C is from the US Department of Energy Genomes Program, http://genomics.energy.gov

For our purposes, the base is the most important portion of the nucleotide. When a gene makes its protein, in order to allow the cell's machinery to read the gene's instructions, the first thing that happens is the bonds between the bases are broken, separating the two strands of DNA, and the DNA double helix opens like a zipper. This produces two single strands of nucleotides, with the bases of the nucleotides exposed, like the teeth of the open zipper. Specialized proteins read the base sequence of one of the two DNA strands (i.e. they read the gene's **coding sequence**), and make a new molecule known as messenger mRNA. The mRNA chain is made by stringing together the appropriate sequence of RNA nucleotides. There are four different nucleotides in mRNA. They use the same A, C and G bases as the DNA nucleotides do, but the mRNA nucleotides use a base called uracil (U) where the DNA nucleotides use the base thymine (T).

When people talk about the "sequence of a gene," they are referring to the sequence of bases in the gene's coding sequence. The specific order of A, C, G and T bases in the gene’s coding sequence determines the specific sequence of amino acids that get chained together to make the protein. The specific sequence of amino acids that get chained together to make the protein determines what function the protein can perform, and what level of activity it will have.

The process whereby the cell uses the DNA's instructions to make the gene's mRNA is complicated, but the important part for you to remember is simple: **the sequence of bases in the gene's coding sequence determines the sequence of bases in the gene's mRNA**. In fact, the sequence of bases in the gene's mRNA is identical to the sequence of bases in the gene's coding sequence, except the mRNA contains Us where the DNA contains Ts. For example, if the gene's coding sequence was ATCCATGCTA, the corresponding mRNA sequence would be AUCCAUGCUA.

If the explanation above sounds complicated, consider that this process is the same process you would go through if you were asked to decorate a room for a party, given a string of lights with empty light sockets, a bucket with four different colored light bulbs in it, and a set of written instructions telling you to put the different colored light bulbs in their sockets in a specific sequence. In the case of the gene making its mRNA, the cell's specialized proteins play the role you would play as the worker who reads the instructions and puts the light bulbs into their sockets in the proper order. The four different nucleotides that are present in the cell are like the four different colored light bulbs in the bucket. Finally, the sequence of bases in the gene's coding strand provides the instructions the cell's specialized proteins use to choose from among the four different nucleotides each time they add a new nucleotide to the mRNA chain, just as your written instructions would tell you which colored light bulb was to be put in each socket.

The Sequence Of Bases In The Gene's mRNA Determines The Sequence Of Amino Acids In The Protein

A protein is constructed by chaining amino acids together. There are 20 different amino acids that are found in human proteins, and your body uses different combinations of them to make the approximately 100,000 different proteins that are in your body. To make a protein, one of the cell's many specialized machines, called a **ribosome**, reads the sequence of bases in the gene's mRNA and uses that as its instructions to chain together the appropriate amino acids.

The process whereby the ribosome chains amino acids together is similar to the process we described above, in which you were asked to decorate a room with a string of colored lights, and given instructions regarding the sequence in which the different colored lights should be strung. This time, however, instead of four different colored light bulbs, you have 20. Apart from that, the process is the same. The 20 amino acids that are present in the cell are like the 20 different colored light bulbs in the bucket. The sequence of bases in the gene's mRNA act like your written instructions, specifying the sequence of amino acids to be chained together. The ribosome chains together the amino acids, just as you would insert the light bulbs into the sockets in the specified order. Some proteins are short (10 amino acids in length), while other proteins are long (thousands of amino acids in length).

The ribosome reads the gene's mRNA three nucleotides at a time. Each set of three nucleotides is called a **codon**. Each codon instructs the ribosome to insert one amino acid into the growing chain. Because there are three nucleotides in a codon, and four possible nucleotides in each of those three positions, there are 64 possible mRNA codons (4 X 4 X 4 = 64). Figure 2 illustrates the **genetic code**, which illustrates the relationship between the base sequence in the mRNA codon and the amino acid that gets incorporated into the protein. As Figure 2 illustrates, when the ribosome reads a CAG codon, it incorporates the amino acid glutamine into the protein. When the ribosome reads an AUU codon, however, it incorporates the amino acid isoleucine into the protein. Notice also that three of the codons are STOP codons. When the ribosome reads a codon with the sequence UGA, UAG or UAA, it stops chaining amino acids together, and releases the chain of amino acids so it can be processed into the protein's final form, to be used by the cell.

First letter	Second letter: U	C	A	G	Third letter
U	UUU UUC Phenylalanine UUA UUG Leucine	UCU UCC UCA UCG Serine	UAU UAC Tyrosine UAA Stop codon UAG Stop codon	UGU UGC Cysteine UGA Stop codon UGG Tryptophan	U C A G
C	CUU CUC CUA CUG Leucine	CCU CCC CCA CCG Proline	CAU CAC Histidine CAA CAG Glutamine	CGU CGC CGA CGG Arginine	U C A G
A	AUU AUC AUA Isoleucine AUG Methionine; start codon	ACU ACC ACA ACG Threonine	AAU AAC Asparagine AAA AAG Lysine	AGU AGC Serine AGA AGG Arginine	U C A G
G	GUU GUC GUA GUG Valine	GCU GCC GCA GCG Alanine	GAU GAC Aspartic acid GAA GAG Glutamic acid	GGU GGC GGA GGG Glycine	U C A G

Figure 2. The genetic code illustrates which amino acids the ribosome incorporates into the protein when it reads each of the 64 different possible mRNA codons.

The Amino Acid Sequence Of A Protein Determines The Function It Can Perform, As Well As The Level Of Activity At Which It Works

Figure 3 below illustrates the structures of the 20 amino acids that make up human proteins. Notice that all the amino acids have a central carbon (C), and amino group (NH_2 or NH_3^+) and a carboxyl group (COOH or COO^-), which is what makes a compound an acid (therefore the name "amino acid").

Note also that they all also have a side chain (also called an organic group, an R group, or a functional group), and that each amino acid's side chain is different from the other amino acids' side chains. Some are larger than others. Some are electrically charged, while others are electrically neutral.

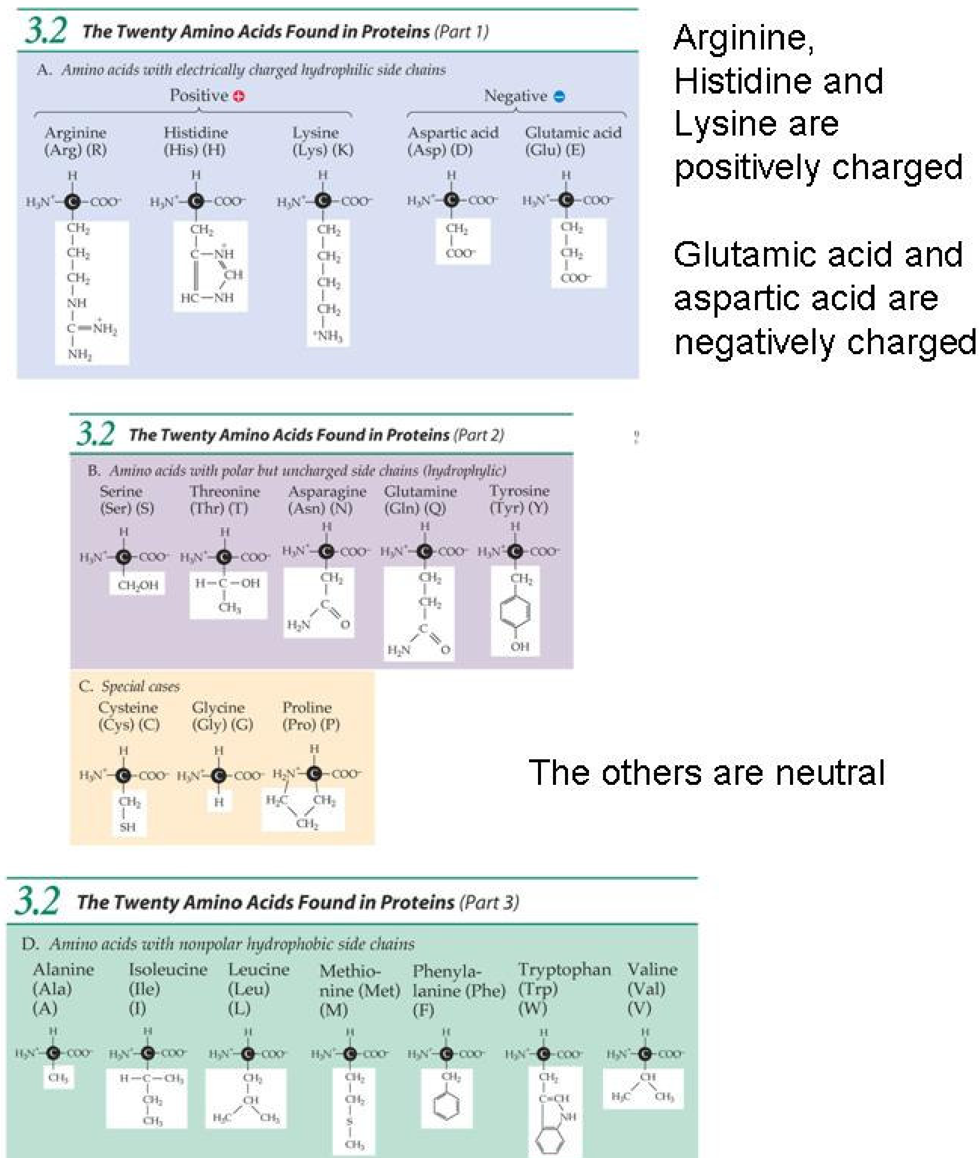

Figure 3. The structures of the 20 amino acids commonly found in human proteins.

In order to perform its function, a protein needs to adopt a specific 3D shape when it is at rest, and it also needs to move to some degree as it performs its function. A protein's 3D shape is maintained in part by attractions and repulsions between positively and negatively charged amino acids. If a protein needs to bend at one point to adopt its 3D shape, it needs to have an amino acid that has a positive charge (ex. arginine) and an amino acid that has a negative charge (ex. glutamic acid) in positions where they can attract each other, causing the protein to bend.

In addition, in order for the protein to move properly, the amino acids must be arranged so that none of them gets in the way of the protein's moving parts. This means that the protein must have amino acids with small side chains (ex. glycine) in places where it needs to bend sharply. If a genetic mutation causes an amino acid with a larger side chain (ex. tyrosine) to be put in place of one with a smaller side chain, having a bulky side chain where you usually have a small one may limit the protein's ability to move the way it needs to in order to perform its function. This may reduce the protein's level of activity by some degree, or it may abolish the protein's activity altogether (we'll discuss this in detail in the next chapter).

With 20 possible amino acids at each position, and virtually no limit to the number of amino acids you can join together to make a protein, you can see that your body has the ability to produce a virtually unlimited number of different proteins. For example, with 20 amino acids to choose from for each position, there are 64,000,000 (64 million) different proteins you can make just by joining together six amino acids (20 X 20 X 20 X 20 X 20 X 20 = 64,000,000). The specific amino acids that get joined together, and the specific order in which they get joined together, determines the 3D shape a protein can adopt, and the degree to which it can move. This in turn determines the function the protein can perform, as well as the level of activity at which it can work.

Chapter Review

- Your proteins perform almost all the functions that are needed to keep you healthy. They digest your food, absorb nutrients, metabolize the drugs you take, repair damaged DNA and detoxify environmental toxins and dangerous by-products of your metabolism, to name just a few of their functions.
- The sequence of bases in the gene's DNA provides the instructions to make the gene's mRNA.
- The sequence of bases in the gene's mRNA tells the cell which specific amino acids to chain together to make that gene's protein.
- The specific sequence of amino acids that get chained together to make the protein determines the 3D shape(s) that the protein can adopt, and also the degree to which the protein can move as it performs its function. These factors in turn determine the function the protein can perform, as well as the level of activity at which the protein works.

Study Questions— The Process Whereby Your Genes Make Your Proteins

1. Please name the three components of a DNA nucleotide. Which is the most important component for our purposes? Why?

2. If the sequence of bases in a portion of a gene's coding sequence was CGTATCGTA, what would the sequence of bases be in the corresponding portion of that gene's mRNA?

3. If Person 1 had the sequence CGG AAG CAG for three codons in Gene A, and Person 2 had the sequence CGG GAG CAG for the same three codons, how would the amino acid sequence in protein A differ between Person 1 and Person 2?