

Bare Bones Biochemistry

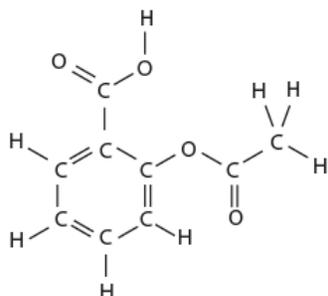
Want to understand this whole synthetic biology thing but don't know a virus from a glucose molecule? This document's for you! To understand synthetic biology, you're going to need to understand some parts of basic organic chemistry, quite a bit of biochemistry, and oodles of molecular biology. Modern biology, or biology as it has been studied since around 1950 or so, is deeply reductionist. By this, I mean it is focused on understanding how living systems work as assemblages of the chemicals from which they are derived. Very broadly, people often refer to this view of biology as "molecular biology," but as you will see, usually this term is used to refer to something much more specific. The distinction I'm making here is the view of biology as classification systems based on macroscopic aspects of things such as asking how many bones are present in your arm, or how many types of plants live in the forest, with what makes these organisms tick at the molecular level. The chemistry and physics of molecules has been grounded in well-tested physical theories since the early 20th century. If we can understand biology in terms of chemistry, then we have good reason to believe that we've learned the full story of how living systems work. Well, that's the goal anyway, and as a devout reductionist you would have to take this for granted as being true. The devil is in the details, and the details of how you go from the world of atoms and energy to a person walking around and thinking gets extremely complicated. If we're going to understand any of this, we need to start by understanding chemistry.

Tinker toy chemistry

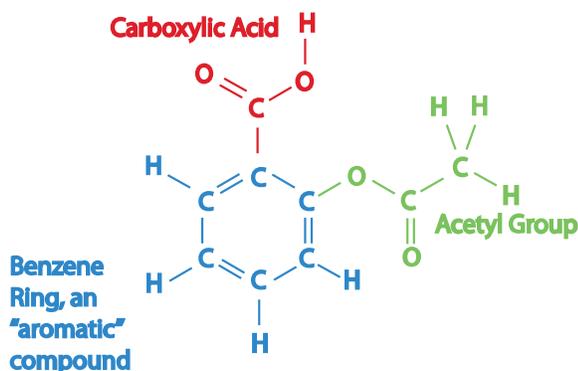
Chemistry is an incredibly deep field. There is wonderful theory that deeply explains our physical world. It's all rooted in quantum mechanics, but fortunately you really don't need to understand any of it to get the basics of biology. Just take me on my word that all this chemistry stuff which I'll get into is rooted in very mathematical and time-tested theories. For our purposes, it is sufficient to think of chemistry as a set of tinker toys. As I'm sure you know, all matter is made up of atoms, and there are 7 major elements involved in biology SCHNOPS—that's sulfur, carbon, hydrogen, nitrogen, oxygen, and phosphorus. There are many others that play important roles such as Ca (calcium), Mg (magnesium), Fe (iron), Na (sodium), K (potassium), Cl (chlorine) and even some important though rare things like Mb (molybdenum) that show up from time to time. However, most of the things we'll be talking about are built out of SCHNOPS and it is sufficient to get the basics of biochemistry if you just understand these elements.

Now, atoms aren't very interesting by themselves. What makes things work is that they interact with each other. There are various types of interactions, but the first one you should understand in the context of biology is the covalent bonds present in "organic" molecules. Groupings of atoms covalently bonded together are referred to as "molecules" and "organic" molecules are the family of compounds that are biologically relevant. By organic, we basically mean carbon-containing compounds. Aspirin, amino acids, sugar, and gasoline are all examples of organic compounds containing carbon. They'll probably revoke my ACS membership for telling you this, but all you really need to know about organic molecules is that they behave like tinker toys. You can think of an atom as a ball with holes drilled into it. A peg sticks into each one of those holes and joins it to another atom. We call those pegs "chemical

bonds.” P makes 5 bonds, C likes 4 bonds, N likes 3 bonds and occasionally 4, S and O like 2 bonds, and H likes 1. To make a “happy” molecule, all of the bonds need to be present. Now, you can make all sorts of things that satisfy the rule that all the bonds are present and still make some very unstable molecules, but the converse is largely true—stable molecules will obey the rule that all the bonds are satisfied. Also, you can have double bonds between atoms. Let’s look at aspirin as an example. It is composed of 3 elements: carbon, hydrogen, and oxygen.

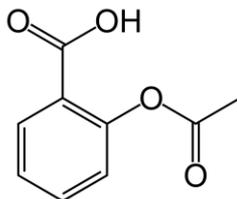


Looking at this structure, examine each atom (represented by its letter code) and count how many sticks are coming out of each. You’ll notice that there is 1 for each H, 4 for each C, and 2 for each O. In some cases, O is attached to two different atoms. In other cases, O has two bonds to a carbon (a double bond). Nevertheless, the main rule is that all the bonds are satisfied which is a prerequisite for it being stable. We have all sorts of names for particular configurations of atoms within organic molecules, and they’ll come up again and again, so it’s worth highlighting a few here:

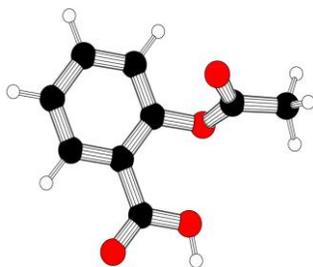


The thing you are looking at is a representation of the chemical structure of an aspirin molecule. Just like the tinker toys, you can apply quite a bit of force on a molecule without changing its structure. It takes some elbow grease to pull the peg out of two tinker toys, and similarly it takes some energy to break the bonds between atoms present in molecules. When the bonds break and change the chemical structure, we call this a chemical reaction. So, chemical reactions in general operate on covalent chemical bonds. Chemical reactions of the sort that happen in biology and in synthetic organic chemistry tend to operate at the level of these “functional groups” and this is why chemists have so many names for the different groupings of atoms. In general, you can’t say much about the behavior of single atoms within molecules, but you can say a lot about the general behavior of functional groups and how they interact with each other and what chemical reactions they’ll engage in.

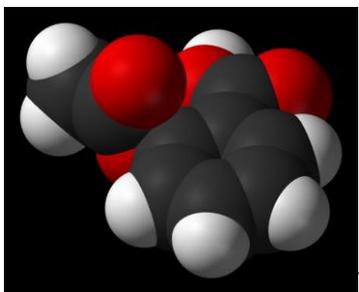
In representing chemical structures, organic chemists rarely if ever draw molecules like I drew aspirin for you. The more typical way of representing the aspirin molecule is:



The reason for this is quite simple. There are carbon atoms all over the place in organic molecules, so you can just assume if you see a corner there is a carbon there. Also, hydrogen is so frequently occupying the remaining bonds on the carbons that there's really no reason to draw them out explicitly. One could look at this simplified structure and mistakenly believe that it isn't obeying the "satisfy all the bonds" rule. Don't be fooled! The "extra" bonds that it needs to look happy are just hydrogens. You'll also often see views of molecules that look like this:

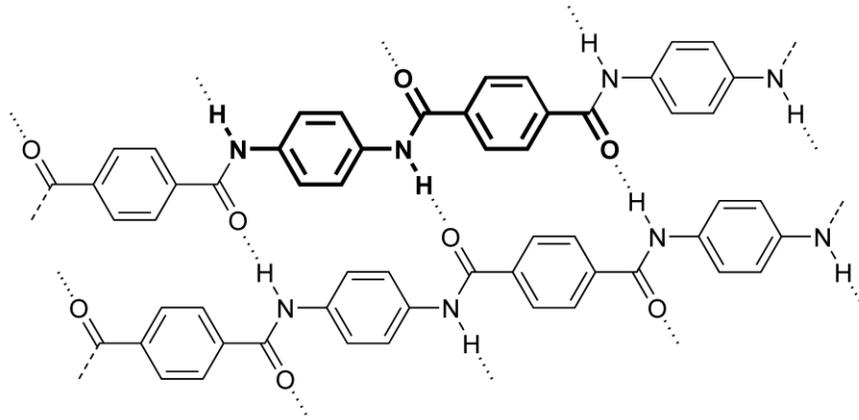


Here, carbon is represented by black balls, hydrogen by white balls, and oxygen by red balls. Usually nitrogen (if present) will show up as blue balls and people get creative with the other atoms. This is a 3D model of the molecule. The center of each ball is present at the relative coordinates of the molecule in space. The length of the bonds are proportional to the true distance between the atoms, and the angles are all how the molecule will naturally bend itself. The reason this molecule will occupy the particular spacial coordinates and angles represented in the 3D model is all deeply explainable with quantum mechanics. For now, don't worry about it—just keep in mind that most molecules aren't really flat as they appeared in the first 2 representations, they have 3D structure to them. The last type of view you'll sometimes see is a "space-filling" model of the molecule. If you had a microscope and could "look" at a molecule, the space filling model is sortov what you would see. Basically, it's the 3D model you saw before, but the real radius of each atom is represented.



Just one word, plastic

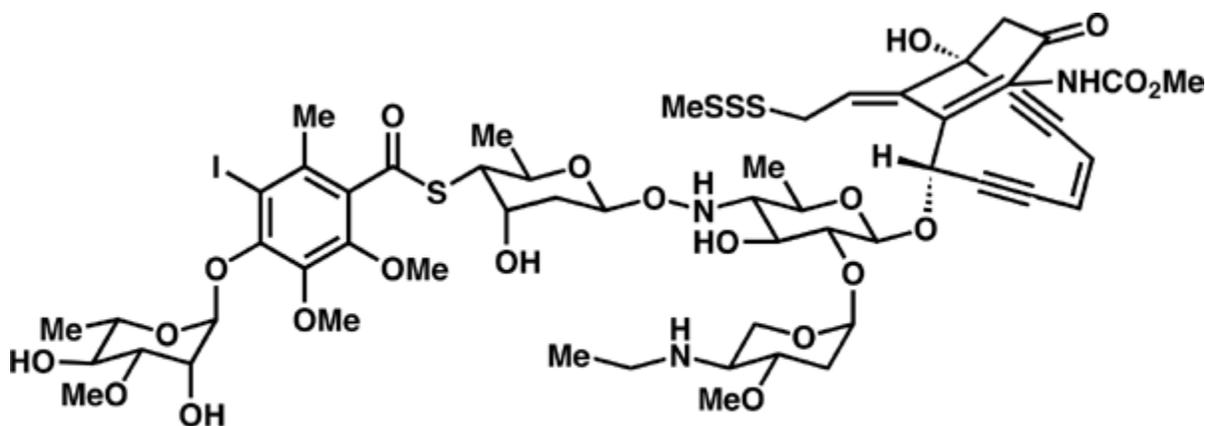
You might be thinking that plastic has nothing to do with how biology works, but you'd be oh so wrong. Well, not plastic per se, but the type of molecules that plastics belong to called "polymers" are major players in biology. Polymers are organic molecules, like aspirin, but they are really large and are made up of repeats of a chemical moieties called "monomers". For example, let's look at Kevlar, a polymer used to make bullet-proof vests:



Do you see the aromatic rings in there just like were present on aspirin? So, it kind of looks like aspirin, but there are two chemical moieties (highlighted in bold) that get repeated over and over. Now, this is actually representing 2 molecules of Kevlar polymer. See the dotted lines? Those aren't covalent bonds—they are "non-covalent bonds" and in particular, they are referred to as "hydrogen bonds". These non-covalent bonds will become super important later for understanding biochemistry, but don't worry about them for now. Just ignore that there are two identical molecules here and focus your eyes on just one of them. Kevlar is an example of a "linear" polymer in that each repeating moiety of its molecules are attached at 2 points to the other moieties. So, it's like pearls on a necklace. The number of repeats of the monomer unit is variable. Here 4 are represented, but in reality polymers usually contain hundreds or thousands of repeats. If you have more than 2 points of attachment between moieties, you'd have a much more complicated polymer as it can be "branched." Most of the molecules we'll be dealing with in biology (and the most important ones) are the linear type, but there are also examples of branched polymers. Now, getting the name polymer usually means that there are many repeats of a common unit. If it is just a few repeats, say 4, you'd typically call the compound an oligomer. We'll have examples of both oligomers and polymers in biology. Another thing to note here is the type of linkage between the monomers. You have "C double bond O" bonded to "NH" which is referred to as an amide group. The amide bond between monomers is going to be really important when we start talking about proteins.

What does this have to do with biology?

All living things -- bacteria, amoebas, humans, plants, mushrooms-- are made up of cells. There are very important differences between the cells in these 5 different examples, but from a chemist's perspective, they are all identical. They are cocktails of water, organic molecules, and salts. You know what water is—that's H₂O, or two hydrogen atoms covalently attached to one oxygen atom (notice it obeys the all bonds satisfied rule!). Salts are things like sodium chloride (NaCl). They are necessary components of cells, but for our purposes they aren't the main players. The ones we care about are the organic molecules. The diversity of organic molecules present in biological systems is the envy of every bench chemist. It is really rich and there are some beautiful looking molecules that nature makes. Like check out this one:



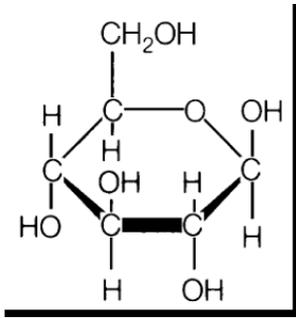
Calicheamicin γ_1

Crazy looking, huh? Cells make some amazing things.

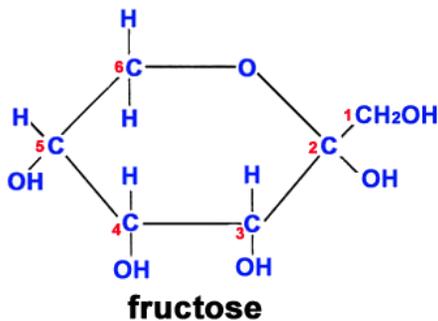
It is useful to lump everything into two categories: primary and secondary metabolism. By this, we mean that there is a list of organic compounds or “metabolites” that are present in every living cell, and there are another set that might be present only in particular species. Drug-like compounds are almost always members of the latter set, but in terms of understanding how cells work, what we really care about are the primary metabolites. We can break these up into 5 categories: **carbohydrates, lipids, amino acids, nucleic acids, and other**. Within each of these categories there are both monomers and polymers. So, let's take a look at what these things are and what role they play in the cell.

Carbohydrates

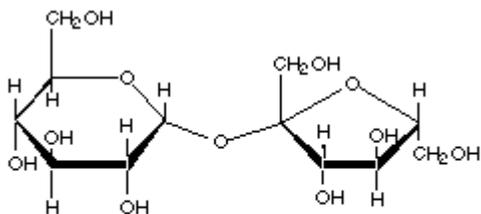
The carbohydrate family of compounds contains monomers and polymers of “monosaccharides” of which the most famous by far is glucose:



Glucose is important because it is a common energy-storing molecule in your body. Heard of “blood glucose levels” in the context of diabetes? That is referring to how much glucose is floating through your bloodstream. Cells can use energy stored in the bonds of glucose. Your body obtains glucose from foods you eat by degrading more complex biochemicals to their monomeric constituents. To illustrate this, let’s look at another monosaccharide, fructose:

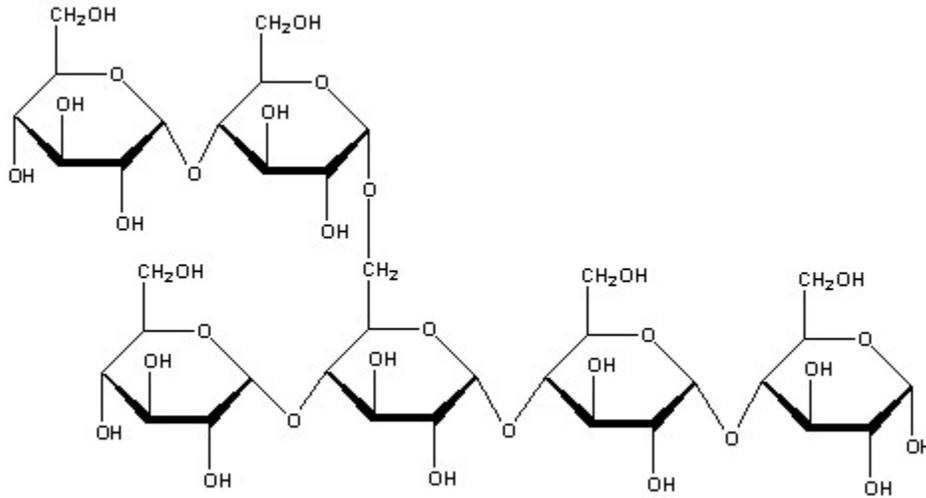


Looks pretty similar, huh. It’s drawn slightly differently, but most of the differences between the monosaccharides are subtle. Monosaccharides are primarily distinguished by having 5 carbons (pentoses), 6 carbons (hexoses), or 7 carbons (heptoses) and whether the hydroxyl groups (which are the -OH’s, an oxygen sandwiched between a carbon and a hydrogen) are pointed up from the ring or down from the ring. Some monosaccharides are amino sugars meaning that one of the hydroxyl groups will be replaced by an amino group (-NH₂, or a nitrogen atom bonded to 2 hydrogens and one carbon). These monosaccharides can exist as single moieties as shown, or they can be disaccharides, oligosaccharides, or polysaccharides which just means a polymer of monosaccharide monomers. So, your body converts polysaccharides and other carbohydrates into monosaccharides in your digestive tract, and these get transported into your bloodstream and made available to cells in your body to use as food. Consider table sugar, or sucrose:

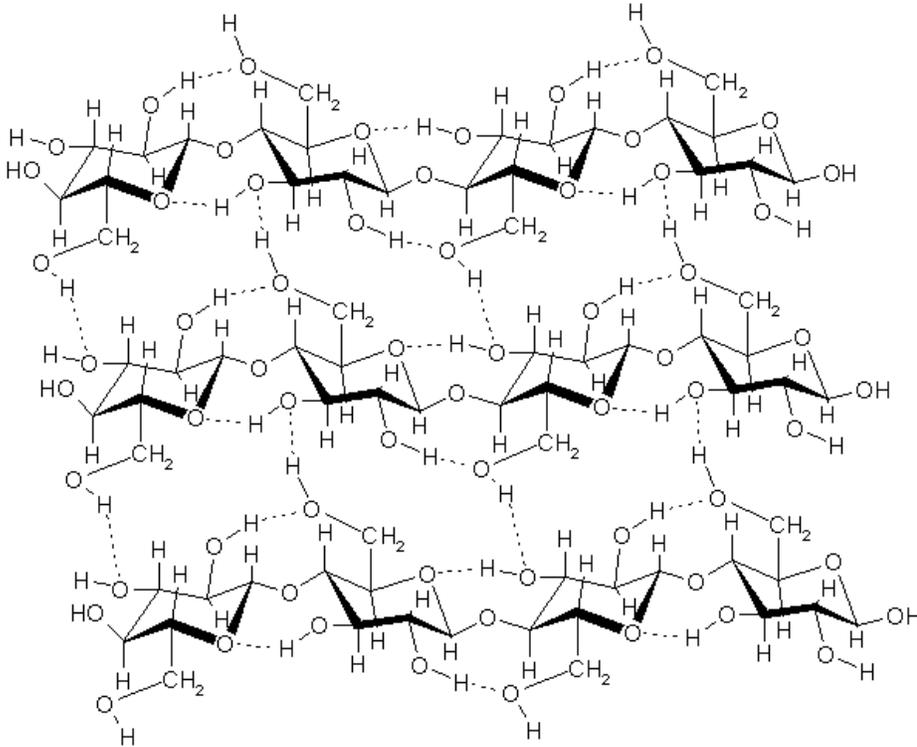


Sucrose
(glucose (α.1-->2) fructose)

It is a glucose covalently linked to a fructose moiety. Your body will break the bond between these two moieties (referred to generally as a glycosidic bond) through a chemical reaction called hydrolysis. Hydrolysis means to break up a molecule with water. So, your body reacts sucrose with water to generate fructose and glucose which then can be used for energy. Sucrose is an example of a disaccharide. Starch is an example of polysaccharide, or a polymer of glucose monomers:



Notice that one of the glucose moieties has 3 glycosidic bonds to other glucoses. It therefore isn't a linear polymer, it's a branched polymer. Your body stores up glucose in your liver as starch. Plants often store it up as well hence much of the carbohydrate that you eat is starch. Another glucose-derived polymer is cellulose:



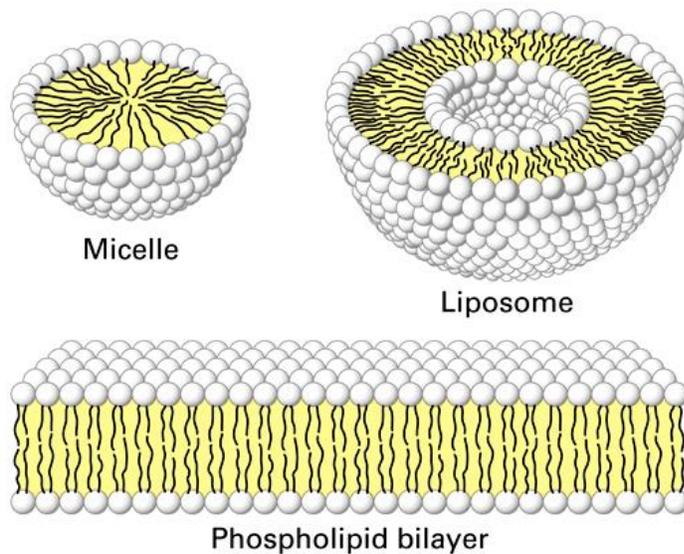
You are looking at 3 chains of linear cellulose polymer chains. See those dotted lines again? Those are again hydrogen bonds. Hydrogen bonds are non-covalent interactions between chemical groups. Many biologically relevant molecules (including water!) make hydrogen bonds with other functional groups. They are weak in strength relative to chemical bonds and are therefore easier to break. They don't obey the "satisfy all the bonds" rule since they create more bonds than these atoms really want. We'll come back to this later, but suffice it to say that one hydrogen bond isn't worth much, but a whole lot of them working together adds up to a lot of interaction energy when the interactions are "cooperative". In cellulose, the way the glucose monomers are attached to one another positions all their hydroxyl groups (the -OH 's) in such a way that they are positioned really well to interact with other cellulose polymer chains by hydrogen bonds. So, you get really long parallel-arranged chains of cellulose all glued together. The result? Paper. Cellulose tends to make for really nicely structured materials that don't dissolve in water, and plants use cellulose to build up their cell walls.

So, carbohydrates play 2 really central roles in biology. First of all, they are very common energy-storage molecules. Secondly, they are often made into structural polymers that help maintain the morphology and physical stress resistance of cells. They are also important in immunology, cancer, and microbiology, but I'm not even going to touch those here. What we're really trying to get to is the "central dogma of molecular biology" which is the core idea of the reductionist view of biology. The carbohydrates don't even show up in the central dogma, so in this sense they aren't very important. There are two monosaccharides that are involved, though, and we'll come back to them in a bit: ribose and deoxyribose.

another to create an ester linkage, and triglycerides have 3 of them. In doing so, they extrude a water molecule (count up the atoms, you'll see that you are missing one O and 2 H's—that becomes the water). Just like glycosidic bonds, you can run that reaction in reverse reacting an ester with water and re-making hydroxyl groups and esters. If you hydrolyze triglycerides, you make old-fashioned soap. Soap makes bubbles and is good for cleaning greasy stuff. Why is this? Well, this gets back into those "non-covalent interactions". In particular, what you are seeing here is the "hydrophobic effect". Notice the long string of carbon chains present on the fatty acid? Think of that as gasoline. Gasoline is principally composed of alkanes also referred to as hydrocarbons. What this means is that they just have carbon and hydrogen. They vary by how many carbons they contain and how they are arranged, but they all look something like this:



If you just replace one of those -CH₃ groups with a -COOH group, you have a fatty acid. Ever tried to mix gasoline and water? The gasoline just floats on top. Hydrocarbons and water do not like to mix. The reason for this is that water likes to hydrogen bond, and hydrocarbons don't have any of the functional groups that can participate in hydrogen bonding. So, hydrocarbons tend to get pushed away from water. This is the hydrophobic effect. One end of the fatty acid chain is a carboxyl group, the other end is hydrocarbon. The carboxylic acid loves to hydrogen bond with water. The other end hates water. The result of this is that if you mix water and fatty acids you'll get something called a micelle, a liposome, or a lipid bilayer:



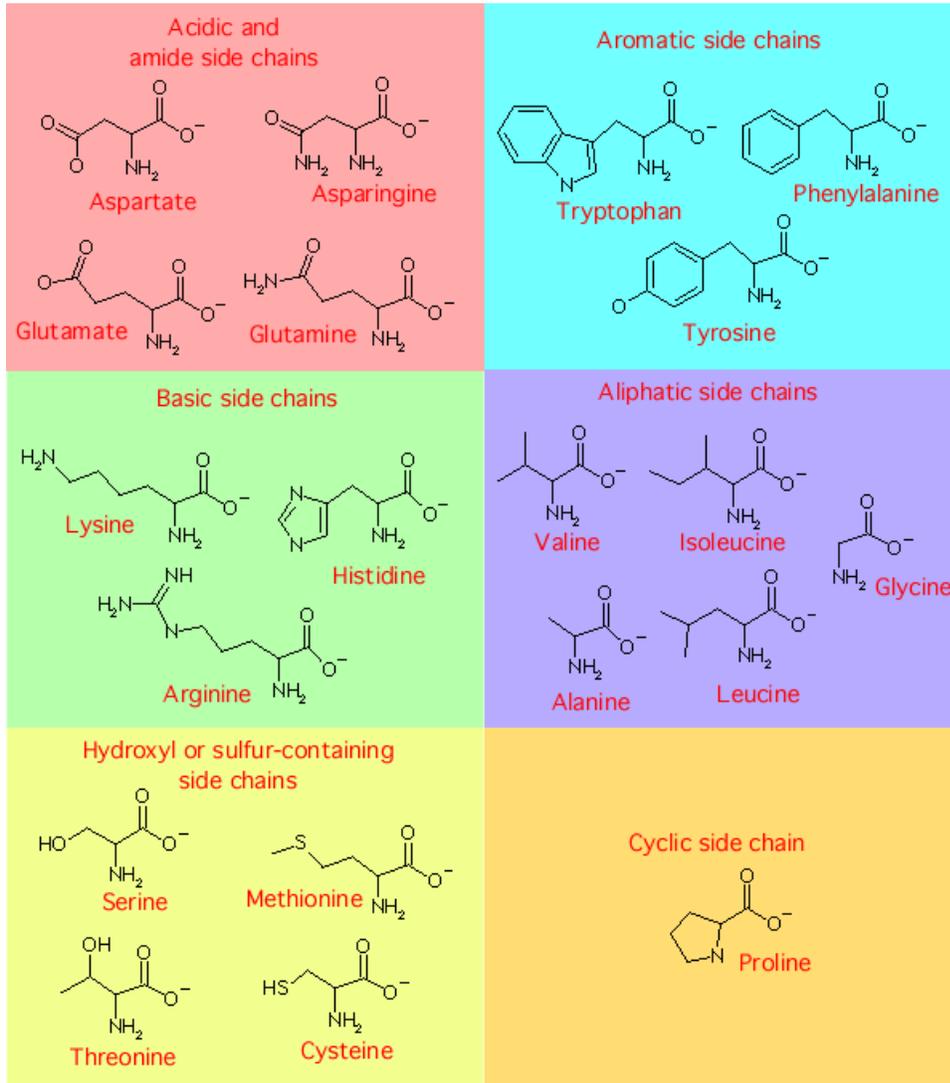
Basically, what you are seeing is the little white balls are the hydrophilic carboxylic acid end of the fatty acid, and the black tails are the hydrocarbon tails. The hydrocarbon tails get pushed towards each other and the carboxylic acid ends move to the outside to be with the water. Different lipids make different types of these structures. With straight-up fatty acids, such as are present in old-fashioned soap, you get micelles. If you mix fatty acids, water, and something greasy like gasoline, the gasoline will get squeezed to the interior of the micelle along with all the fatty acid chains, but the whole thing will float around as a spherical aggregate in water. When you make triglycerides, the resulting lipids tend to form the liposome and bilayer types of structures. These structures present a barrier between two different water based (termed “aqueous”) regions and are the basis for the inside and outside of the cell. Every cell is surrounded by at least one bilayer membrane composed of lipids.

So, the principle place you find lipids in a cell are in the membranes as a structural component. They are a pretty diverse group of chemicals, though. They all have fatty acid chains—that’s the defining characteristic, but how many chains and what they are attached to is diverse. They can also play roles in cell signaling and various specific physiological roles, but we won’t get into that here. Again, we’re trying to focus on the central dogma, and well, lipids just aren’t a part of it. So, let’s not worry any more about these.

Amino acids, peptides, and proteins

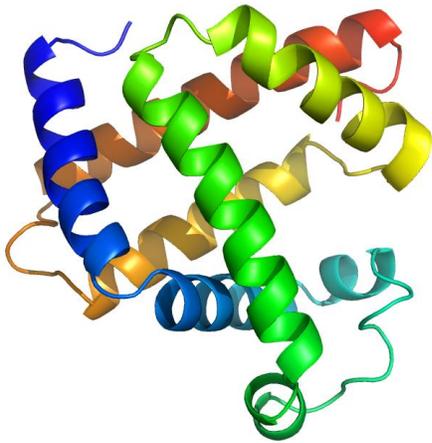
Alright, now we’re getting to key players in the central dogma—the proteins. Proteins are linear polymers of typically 100-1000 repeats of amino acids. Now, we’ll say a lot more about proteins as they are the worker molecules in the cell—they do the magic that happens in cells. For now, let’s just focus on the amino acids and their chemical properties. There are 20 amino acids to choose from, and the

ordering of these monomer units defines the sequence and properties of a protein. Here are the amino acids:



You'll notice that all of them have an amino group (-NH₂) and a carboxyl group (-COOH). The rest of the structure is referred to as the "side chain" and it is the feature that distinguishes the amino acids from one another. To make a polymer of amino acids, you link them by reacting the amino group of one to the carboxyl group of the other and extrude a water molecule. The result is an "amide bonds" and is referred to more specifically as a "peptide bond". Note that Kevlar, the plastic we saw earlier, was also linked together by amide bonds. The distinction between a "peptide" and a "protein" is really subtle. All proteins are peptides—peptide means any molecule that is a polymer of amino acids. A protein implies that the thing is pretty big (say, over 50 amino acids at least, usually much larger) and also that it has some tertiary structure associated with it. Proteins always have a defined sequence. By this we mean that the order and sequence of the amino acids in the peptide chain are the same in every molecule. This is referred to as its "primary structure"

Enzymes, the “twitching” molecule in muscles, the receptors that detect chemicals when you smell something, the structural material of your skin, and the light receptors that allow you to see things are all at their core protein functions. The function of a protein is somehow encapsulated in its amino acids sequence, but just how they do that gets really complicated. At the heart of it is protein folding. If you just take an arbitrary protein sequence, it just floats around in solution as a floppy molecule and does nothing special. Protein sequence space is something like 20^{500} , and a rare subset of those sequences have the special ability that they can fold into a stable structure. You can get local folding into things called alpha-helices and beta-strands, and that is called secondary structure. However, you can get more elaborate tertiary structures as well. Here is an example of a tertiary structure of a protein called myoglobin which binds to oxygen in muscle cells:

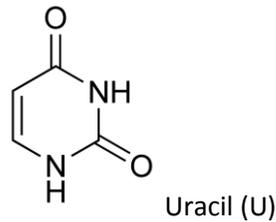
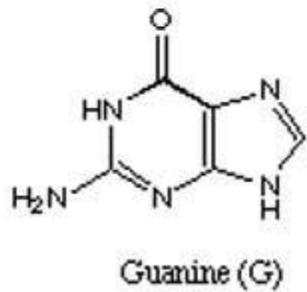
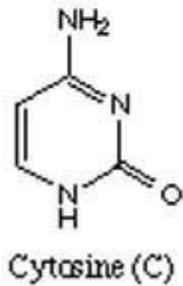
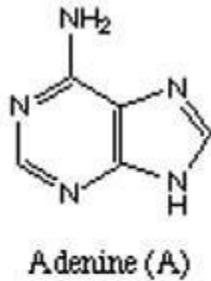
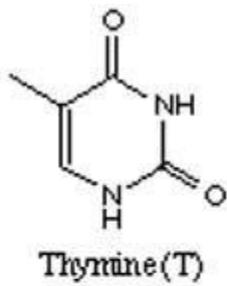


Notice the coil things? Those are alpha helices. Now, how this folding business happens has everything to do with the full arsenal of non-covalent interactions. At the end of it, though, you get a fairly rigid ball of functional groups after folding. These structures can then interact with other biochemicals, bind to them, catalyze chemical reactions with them, and so on.

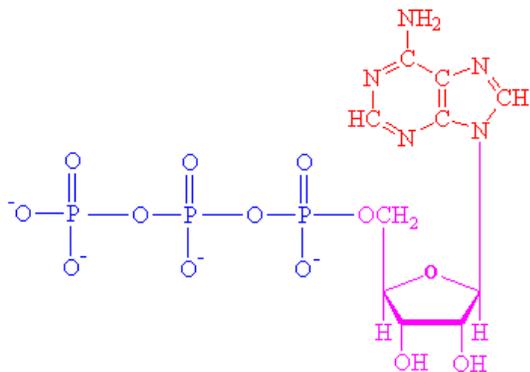
Nucleic Acids

The other major players in the central dogma are the nucleic acids RNA and DNA. First though, let's start with the defining components of the nucleic acid family: the nucleobases adenine, thymine, uracil, cytosine, and guanine.

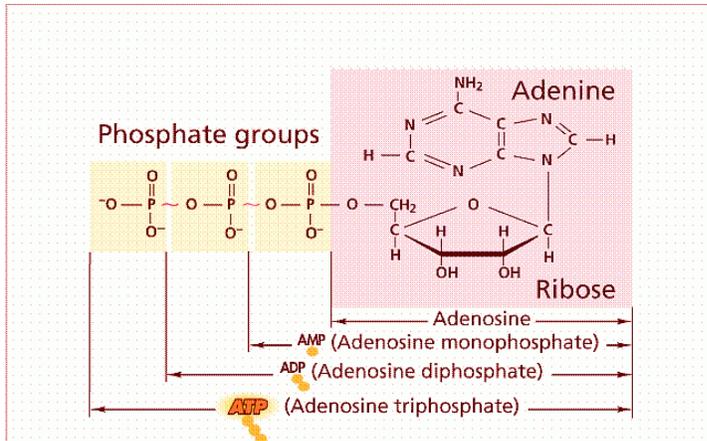
The nucleobases



Everything in the nucleic acid family contain at least one of these bases. The simplest biologically meaningful nucleic acid molecules are the nucleotides, the most important of which is adenosine triphosphate (ATP):

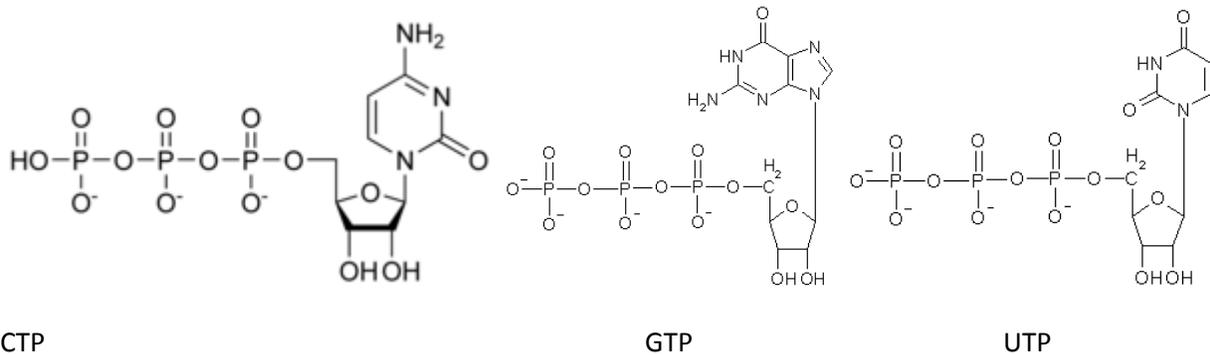


There are two moieties of the ATP molecule that you have seen before. In red, you see Adenine, and in pink you have the monosaccharide ribose. You also have 3 “phosphates” attached to the molecule. ATP is the primary intracellular energy molecule. We noted before that glucose is an important energy molecule in the bloodstream. When cells (all kinds of cell—bacteria, human cells, yeasts) see glucose in their environment, they take it into the cell and do a process called glycolysis and respiration to make ATP. So, outside the cell glucose is a common energy currency. Inside the cell ATP is the major energy currency. When a cell “makes ATP” what we really mean is it adds phosphates to shorter molecules. The total number of ATP, AMP, and ADP molecules in the cell is pretty much constant:

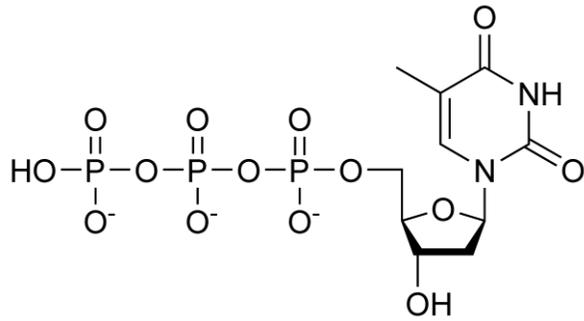


All sorts of processes in the cell will break off the “third” phosphate of ATP to liberate phosphate and ADP and use the liberated energy to drive chemical reactions. Glycolysis and respiration put the phosphate back on.

The other nucleic acids are all based on ATP-like molecules derived from the other bases using either ribose or deoxyribose. Here are the nucleotides based on ribose:



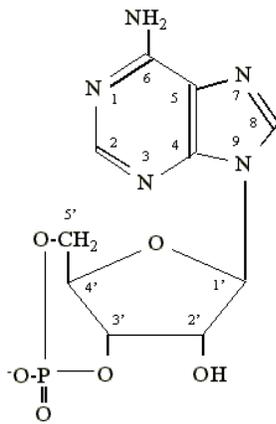
The deoxyribose-based nucleotides are called dATP, dTTP, dCTP, and dGTP. The difference between the ribose-derived and deoxyribose-derived nucleotides is the presence or absence of a “2'-hydroxyl” group on the monosaccharide moiety. Look at the structures of ribose and deoxyribose again. The 2 in “2'-hydroxyl” refers to the 2nd carbon after the ring oxygen. So, start on the oxygen and count 2 carbons over. With an –OH there (the hydroxyl) you have ribose. With an –H there, you have deoxyribose. The other difference between the NTPs and the dNTPs is that the ribose-derived ones use A, C, G, and U (uridine) while the dNTPs use A, C, G, and T (thymine). The difference is just a methyl group (-CH₃), so it's a little subtle. T and U are pretty darn similar bases.



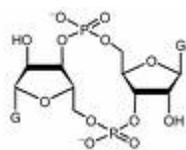
dTTP

Other small molecule nucleic acids

As small molecules, the monophosphate, diphosphate, and triphosphate variants of all these nucleotides are all present and being loaded and unloaded with phosphate. Cyclic and dimeric versions of these compounds have important roles as chemical signals as “secondary messengers” and “cell-cell signaling” molecules. The most famous of these are cAMP and c-di-GMP:

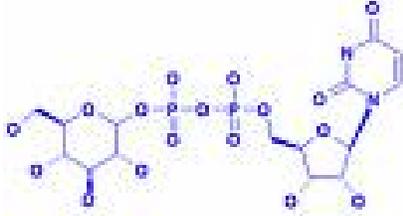


cyclic AMP



cyclic-di-GMP

Nucleotides are also often parts of larger molecules used in biosynthesis. For example, when cells build carbohydrate polymers, they “charge” the monosaccharide by reacting it with various NTPs, the most common of which is UDP:



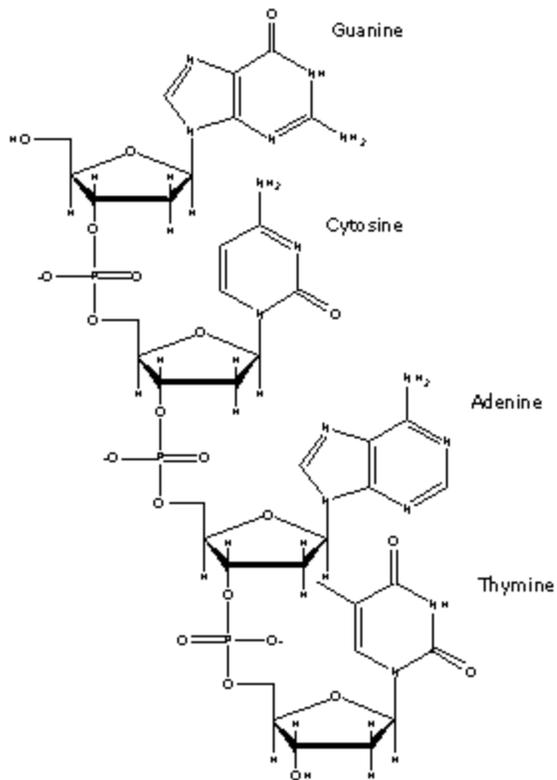
UDP-glucose

When cells make starch from glucose, they first charge up the glucose by reacting it with ATP to generate UDP-glucose. In so doing, they release a phosphate and that liberates energy driving the reaction to completion. The glucose is then transferred from UDP glucose to the growing starch chain releasing UDP. The UDP hydrolysis event also releases energy. The cell then “reloads” the UDP back to UTP by reacting it with ATP. The ATP is of course being continuously generated by glycolysis and respiration. So, the net effect is that the reaction making starch is driven to completion by paying for it in the form of the common energy currency of the cell, ATP.

In our goal of getting to the central dogma, the most important nucleic acids are the polymers of dNTPs (DNA) and NTPs (RNA).

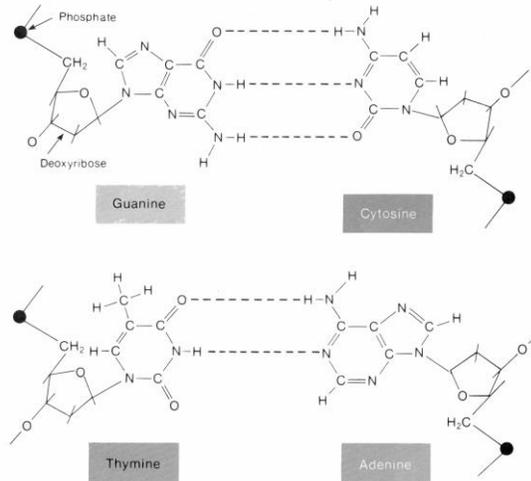
DNA (deoxyribonucleic acid)

If you polymerize a bunch of dNTPs together by reacting the 3' hydroxyl of one dNTP with the alpha phosphate of another dNTP, you make a DNA molecule. Here is an example of a DNA molecule composed of 4 nucleotides:



The “sequence” of this molecule would be GCAT. Why not TACG you ask? There is directionality to a DNA molecule. The “G” is at the 5’ end of the molecule (notice that the 5’ hydroxyl of “G” isn’t bonded to anything else) and the “T” is the 3’ end (notice that the 3’-OH isn’t bonded to anything, but the 5’ hydroxyl of “T” is bonded to a phosphate). By convention, we always write out the sequence of a DNA in the 5’ to 3’ direction. So, GCAT and TACG are not the same molecule.

If you have a short DNA like this and dissolve it in water, it will exist as a single-stranded, linear DNA. By linear, we mean that it has free 5’ and 3’ ends. In contrast, a circular DNA would have its 5’ and 3’ ends linked to one another, so there is no distinguishing feature to determine the “ends” of a circular DNA. A circular DNA still has directionality, though. By single-stranded, we are referring to the wonderful property of nucleic acids that makes this whole life thing work: DNAs like to hydrogen bond with other DNAs, but they do so in a sequence-specific way according to simple rules.



Remember the hydrogen bonding, non-covalent interaction business? Well, A's and T's hydrogen bond with each other. C's and G's hydrogen bond with each other. The other combination don't do anything. In terms of the energetic of the interaction, the base-pairing of two DNAs is highly cooperative. To understand cooperativity, let's do a thought experiment:

Imagine you have a yard stick and you glue a small magnet at the 1 ft mark. You have a second yard stick with a piece of iron at the 1 ft mark. Now, hold the two yard sticks right next to each other, and you find that they stick to each other a little bit. Now, add more magnets—put a small magnet at every inch mark and a piece of iron at every inch mark on the other. Now the two sticks really like to stick to one another. Now, take certain pairs of magnet and iron pieces and switch them on the two sticks. The sticks should attract each other just as before. The binary pattern of magnet-iron-iron-magnet-iron-magnet-magnet-magnet-magnet-iron-magnet can be thought of its sequence. If you have the complement of that sequence on the other stick, the two things stick to one another. OK, same deal with the DNAs. If you have at least 15 bases in a row on a DNA strand, it will base pair under physiologically-meaningful conditions with another DNA strand of the (reverse) complementary sequence. That “reverse” business refers to the fact that for structural regions two nucleic acid strands have to be antiparallel to one another for the base pairing to work. So, it's as if you only can make your yard stick attract another yardstick that's been flipped over.

So, if both a DNA molecule and its reverse complement coexist in the same vicinity, they will base-pair with each other resulting in a double-stranded DNA. In practice, DNAs in cells are always double stranded DNAs. Single-stranded DNAs transiently exist within cells but, stable DNAs are always made as double stranded entities. Cellular DNAs can be linear, circular, or capped by proteins depending on the organism and situation.

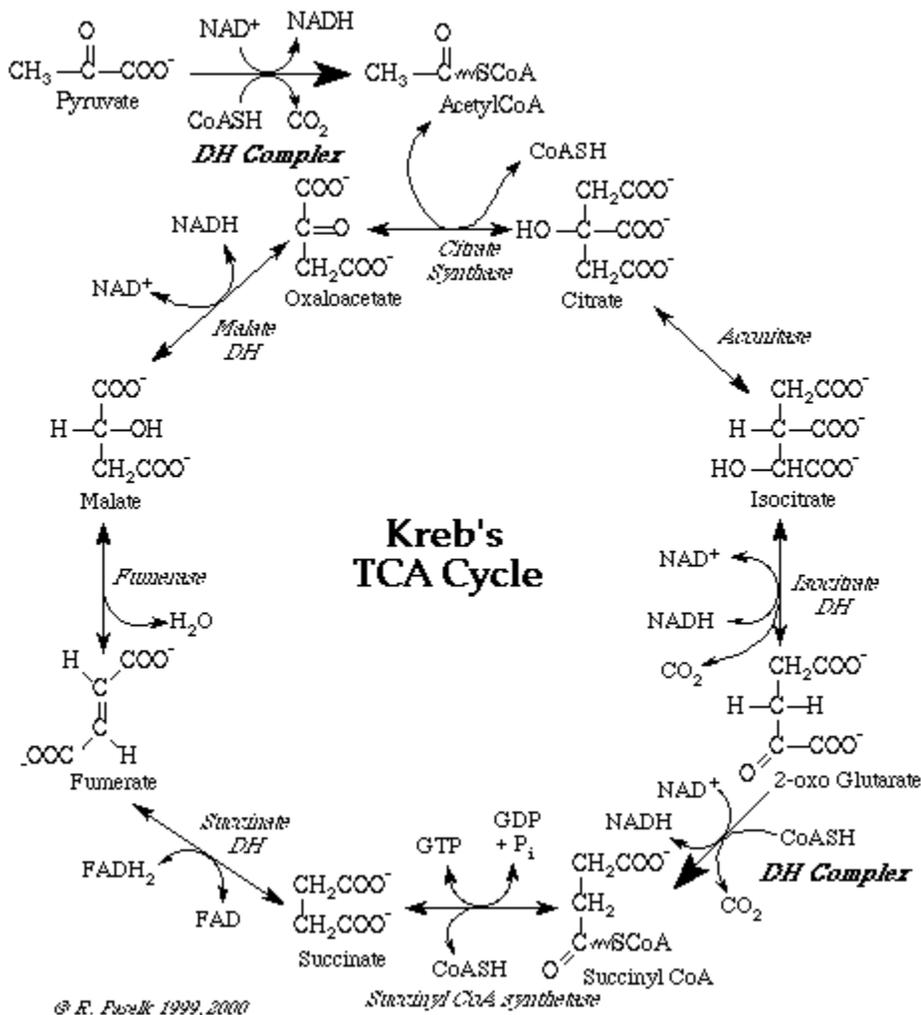
RNA (ribonucleic acid)

The structural difference between DNA and RNA is the 2' hydroxyl group and the use of UTP instead of TTP. Other than that, you polymerize ribonucleotides 5' to 3' just like DNA and they have pretty similar chemical properties. RNA can do the base-pairing thing just like DNA, but typically the reverse-

complementary strand of an RNA is not made within cells. So, they exist principally as single-stranded molecules. This does not mean that RNAs don't base pair in the context of a cell. Regions of an RNA molecule frequently interact with complementary regions of other RNA molecules. Also, RNA molecules can base-pair with themselves resulting in folded structures that sometimes even have catalytic activity.

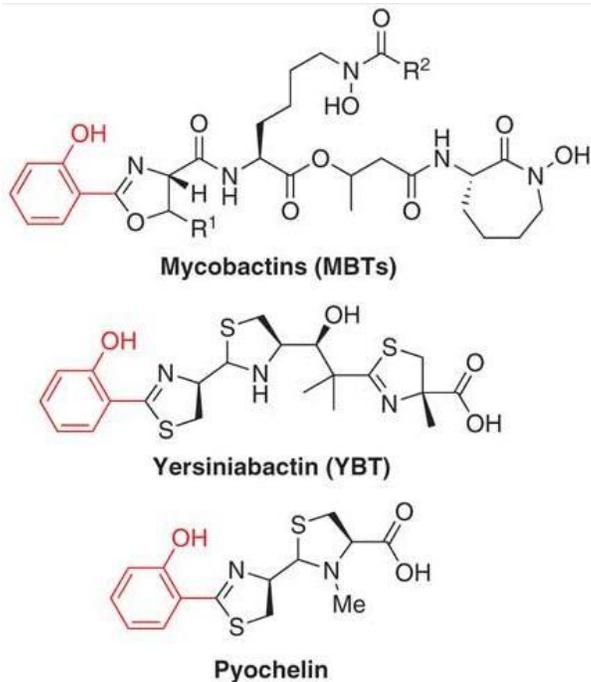
Other

The "other" primary metabolites are all the various things that show up on charts. Most of them are small molecules rather than polymers. In general, they may play signaling roles in cells, but in general they are not information-bearing molecules. Most of them are either intermediates in the pathways that lead to the amino acids, lipids, carbohydrates, and so forth or are intermediates in the breakdown of these chemicals. One important catabolic pathway (catabolic means "breaking down") involved in consuming glucose is the TCA cycle:



So, you have multiple biochemical intermediates, and each arrow represents a distinct chemical reaction that converts one metabolite into another. Each one of these reactions would be catalyzed by one or more enzymes, which are a functional category of proteins.

Secondary metabolites are things made by a cell that aren't explicitly needed for the primary functioning of the cell. They may nevertheless be very important in the context of the organism's lifestyle. For example, many pathogenic bacteria will make small molecules called siderophores that they use to scavenge iron from animal body fluids:



In the context of growing in the lab, they don't need to make the siderophores, and they are entirely specific to species of bacteria. For example, only bacteria in the genus "*Yersinia*" will make Yersiniabactin. *E. coli* makes one called enterochelin despite the fact that *Escherichia* and *Yersinia* are highly similar genera. In the context of causing infections, they are absolutely critical to the survival of the organism. So, there are all sorts of things that cells make that aren't part of primary metabolism but are nonetheless critical to their survival.

Is that it?

Not at all. I've grossly oversimplified this to get you to the basics of what's in a cell and what are the key players in the Central Dogma. You can proceed to the "Basic Molecular Biology" page now and learn about the Central Dogma. If you want to learn more, the Biochemistry text by Garrett and Grisham is fantastic for learning more about primary metabolism. If you want a more traditional full-scope biochemistry text, check out Stryer's Biochemistry textbook. It gets into a lot more protein structure and function stuff and also the physical chemistry of biological systems.