

MIT OpenCourseWare  
<http://ocw.mit.edu>

7.014 Introductory Biology, Spring 2005

Please use the following citation format:

Penny Chisholm, Graham Walker, Julia Khodor, and Michelle Mischke, *7.014 Introductory Biology, Spring 2005*. (Massachusetts Institute of Technology: MIT OpenCourseWare). <http://ocw.mit.edu> (accessed MM DD, YYYY). License: Creative Commons Attribution-Noncommercial-Share Alike.

Note: Please use the actual date you accessed this material in your citation.

For more information about citing these materials or our Terms of Use, visit:  
<http://ocw.mit.edu/terms>

MIT OpenCourseWare  
<http://ocw.mit.edu>

7.014 Introductory Biology, Spring 2005  
Transcript – Lecture 3

So our next class of biomolecule that we're going to talk about are nucleic acids. And we can, for the most part, describe their properties by considering just covalent bonds and hydrogen bonds. Although, that's a bit of an oversimplification. But, anyway, these are, again, polymers.

So this is DNA and RNA, terms you've undoubtedly heard. And these are made by splitting out water. And, in this case, the monomeric units are given the special term nucleotide. And a nucleotide consists of a sugar with something called a base on it.

It's got a phosphate group at one end and a hydroxyl group, one of the sugar hydroxyls that we saw the other day at the other end.

The B stands for base.

And the way the bond is formed, as I said, is by splitting out water like that to form what's known as a phosphodiester bond. And we'll be talking a lot about those when we talk about DNA and RNA in more detail later on in the course. The sugars are pentoses where N equals 5. We were talking about these the other day. The base goes in this position. That's the 1 position of the carbon.

This is the 5 position of the carbon. And this is where the phosphate is located. This sugar is called ribose. And then RNA, which is the polymer of nucleotides that have ribose as the sugar is ribonucleic acid or RNA, as you've known it. If this hydroxyl here is replaced by a hydrogen and the rest of it's the same --

-- this is deoxyribonucleoside. And if you polymerize that together then you get DNA or deoxyribonucleic acid.

The bases come in two flavors. And this will be on your handout. Ones that either have two rings, adenine or guanine. And the general term of those is purine or they have one ring of pyrimidine. And in DNA one finds cytosine and thiamine abbreviated as C and T. Or in RNA, instead of finding thiamine you find uracil, which is the same except that it doesn't have the methyl group that's present at this position on thiamine.

And the important thing about these particular nucleotide bases is that they can form hydrogen bonds in a very special way. It's diagramed on here that this is a guanine pairing with the cytosine, so a G pairing with a C. And you can form

three hydrogen bonds. Or between an A, an adenine and a thiamine you can form two hydrogen bonds.

Those are just the things we were diagramming on the board the other day. And those are the forces that hold the strands of DNA together so that DNA is the double helix, as you know. It's basically a backbone with sugars and phosphates. And then there'll be some sequence of bases down this. And then on the other strand you'll have the base that can form hydrogen bonds with this.

So C, there would be three hydrogen bonds here. This would be a G on this side, again, three hydrogen bonds. If there's an A here that will be a T there, two hydrogen bonds. And so on down. And we'll talk about the implications of this later in the course when we talk about DNA replication, but for the moment I think your eye can see, you can probably see that the geometric arrangement of these is just exactly the same, whether it's a G-C or it's an A-T base pair. You can superimpose them, and they have just exactly the same molecular structure.

And that's really crucial for a lot of things having to do with DNA. So, as you know, it's not just sort of a ladder with hydrogen bonds. It's twisted in 3-dimensional space. That's the double helix. And in that little movie I showed you the nitrogen atoms and the bases are blue so you can pretty much pick it out that there's a series of hydrogen bonds going right down the middle of a DNA molecule with the phosphoribose backbone on the outside.

So every one of your cells, since you have about 3 billion base pairs you have two point something times that many hydrogen bonds holding your DNA together. The thing to remember about the strength of the hydrogen bond, it's about a twentieth a covalent bond, and so you're able to pull those things apart and then put them back together at physiological temperature while leaving all the covalent bonds that make up each strand of the DNA leaving those intact.

It's also possible, since RNAs are usually single-stranded, that if you have a little sequence here that has the complimentary sequence over there then these can pair like this forming a hairpin or some little structure like that. And, again, we'll talk about transfer RNAs which play a really key role in protein synthesis. They're the little translators that go back and forth between the nucleic acid code, the genetic code and the protein code, which is written in amino acids.

And this just shows making an RNA copy for a tRNA gene from the DNA, but then these are the relationships between the complimentary sequences right in that strand. So this thing is able to fold up into a sort of cloverleaf structure that some of you have certainly probably seen at some point. It's a little bit twisted here because you can see how the complimentary sequences have found each other.

And even though this is just a single strand of RNA, by forming hydrogen bonds to complimentary sequences within itself it can take up a structure. And I'll show you. It actually goes on, there are some other forces that come in. And this will fold up into a 3-dimensional structure that goes even beyond what I've shown you, but we won't need to talk about that for just a little bit. OK. So then the next

--

-- class of molecules that we're going to spend a lot of this course on are proteins. And these are polymers again.

-- made by splitting out water. So that's been true of polysaccharides. It's true of nucleic acids. It's true of proteins. In this case the monomers are structures known as amino acids.

And they have an amino group. And then it's joined to a carbon known as the alpha carbon. And then there's a carboxyl group. So this is why they're called amino acids, because their carboxyl group is an acid.

And the way they form --

We'll give these different side chains here. I'll tell you about these side chains in just a minute. The way these form a bond is by splitting out water here. And then this will give this very important bond in nature --

-- which is known as the peptide bond. And there's a chemical property of this that's important. Someone was bemoaning the fact that I had to go over a bunch of chemistry and they hadn't liked 5.011. My apologies.

But we won't be spending all course doing chemistry. But if you want to understand how these things work you do need to understand some of the chemical principles to understand them. And this is a case where it's really important because, although it's written this way with the double bond here and a single bond there, this double bond actually sort of spends part of its time over here. So this is actually sort of a partial double bond. And that has an important consequence because if you're a single bond, if you remember a single bond can bend and stretch but it can also rotate.

But if you're a double bond you cannot rotate. So the peptide bond, and you make a lot of these when you're polymerizing amino acids together to make proteins, those bonds have a very special character that they cannot rotate. Now, let me say, I'll come back and show you why that's important in just a moment. But let me just say a word about the side chains. There are 20 different amino acids.

And they have side chains that have very different chemical properties. And when we start thinking about how a chain of amino acids take up the properties

that make it into an enzyme or part of a motor or a structural protein or into your finger nails or your hair or skin, they have to have very special properties. And it's the sequence of these different amino acids with their different chemical properties that are eventually going to let each protein form up to one particular 3-dimensional structure that will give it its characteristics.

So the different types of amino acids, and again you won't have to memorize these, but here they are up here. But let me just point out the important classes, because the thing you really want to do with this one is to remember the types of amino acids we find. There are negatively charged amino acids.

An example of this would be aspartate. Under physiological conditions, although this is an acid, it will dissociate so it will have a negative charge. And that's abbreviated as A-S-P. Glutamine is another one that has a negative charge. There are also amino acids that have positive charges on the side chain.

An example of this would be lysine which has four methylene groups. And then it has an amino group. But, again, under physiological conditions, around pH 7, that will be protonated so it will have a plus charge. And that is lysine or L-Y-S. And arginine and histidine are two other amino acids that have a positive charge, or can have a positive charge.

Then there's a set of amino acids that have a polar character. They don't have a full charge. And, as you might guess, they have one of the bonds that we've talked about that are polar. This is serine.

Serine. There's another one that has a hydroxyl that's known as threonine.

And then there is a glutamine and asparagine, both of which have an N-H bond. So just through what I've told you here, we haven't even been through the set, you can see how you can begin to decorate an amino acid chain. So there's a plus charge and a minus charge, a polar charge. There's a tremendous amount of diversity because at every single thing you have a choice of 20 things you can put in. So they not only have size and shape characteristics, but they have particular charges and other properties.

Then, as always, there are a bunch of special, oh, excuse me. Actually, before we do that, we have hydrophobic.

Or you could think of these as greasy or water-hating. These are the ones that are sort of when I was talking about trying to dissolve butter into water.

These are things that don't like to interact with water or cannot interact with water, and so they cannot form hydrogen bonds so you cannot get them to go under water easily. And they come from very simple ones that have just the methyl group which is alanine or A-L-A or one like this which would be CH<sub>2</sub>-CH

with a couple of methyl groups. This is even more water-hating, that would be leucine, L-E-U.

Or here's one that you probably could guess that really doesn't interact with water. This is phenylalanine or P-H-E. And you can see what this side chain is. It's a methylene group. And what's dangling off it but a benzene ring. And I think most of you remember from probably beginning chemistry that benzene is something that you cannot dissolve sugar in or something.

It's an organic solvent. It will only dissolve things that have a very hydrophobic character. Then there are some special cases.

Glycine is one, because in this case the side chain is simply a hydrogen atom. And, as a consequence to that, this is a very flexible amino acid. So if you want to build --

If nature wants to build a loop into a protein, it's going to undergo a tight turn. You often find glycines there because there's not a big side chain to get in the way if you're going to be bending the chain in 3-dimensional space. Another one is cysteine, which looks like serine over there, but it has a thiol group instead of a hydroxyl group.

And that's important because that allows for the formation of another special type of bond that if you have one chain of protein that has a cysteine on it and another polypeptide chain that has a cysteine on it and they're close together in space, what can happen is you can form a covalent bond between these under oxidative conditions. This is known as a disulfide bond.

That's covalent. So those two chains, if that bond occurs, are now sort of semi-permanently locked together. They're locked together in a very, very strong way. So this is a feature of, this is the only intrastrand covalent bond that you'd characteristically find in proteins. All the rest of them we're going to show you, when they fold up in 3-dimensional space, depend on other kinds of interactions.

And finally there's one last case which is proline. And this one is a little different because in the amino acid the side chain bends around like this and joins here. So it's actually forming a little circle here between the nitrogen, the amino group and the carboxyl group.

And the consequence of this is this bond cannot rotate. The bond that would normally be able to rotate is not able to do that. And so this is sort of a protein you find that when there are some of these regular structures, I'm going to show you in a minute, like helices and things, this protein won't, this amino acid particularly won't fit into those structures. So you tend to, if nature wants to interrupt a particular regular structure that's coming, it will often find a proline right at that particular point.

OK. So what we've talked about up until now is sort of just the very, very basic piece of protein structure.

It's what called the primary structure which is nothing more than the sequence of amino acids. However, here's a little piece of protein. This is polyalanine. And one thing you can sort of see is if I was trying to figure out how to fold this up into a 3-dimensional confirmation.

And let's say this had 300 amino acids or something, there are essentially an infinite number of confirmations. And so one of the real holy grails still in biology is trying to understand if you see the linear sequence of an amino acid, which we can now deduce, excuse me, of a protein, of an amino. If you see the linear sequence of amino acids in a protein, and we can deduce those from analyzing genomes and so on, how do we go from a thing that says a tryptophan, a cysteine, a serine, a serine, a threonine, whatever down the chain to finding its 3-dimensional structure and ultimately its role?

And you can sort of hopefully get a sense from this of why it's important. So there are levels this goes. The next level is what's known as secondary structure. These are regions of local secondary structure and they're determined by hydrogen bonds.

And I'll show you how these go in just a second. And then you can think about proteins in the tertiary structure. So what we've done and sort of taken a chain and then found out how a little region might take up a particular, for example, here's a portion that's in a helix. This is fairly rigid right now because of the way it's held together. And we'll then find maybe another region like a beta sheet I'm going to show you in a minute.

Ultimately we have to figure out how all these units fold up into a 3-dimensional structure. And what we get there is called the tertiary structure. And this has some other forces we're going to talk about besides covalent bonds and hydrogen bonds that determine that. And then, as I've tried to tell you, you can see that proteins play a lot of roles in nature and they're not all single proteins running around being an enzyme or something like that.

Many of them are parts of machines so they're made to fit together in absolutely beautiful ways. Some of them have, at this point, fifty-hundred parts that all go together fitting shapes and interacting with these shapes on the principles that we'll be talking about here, the different forces that make things happen in nature. And so quaternary means the structure when there's more than one polypeptide chain.

So getting a handle on protein structure was kind of a very important intractable problem for a long time because it was just too hard a nut to crack, but in the

1930s and 1940s x-ray crystallography started to come into usage where basically you'd bounce x-rays off of a crystal.

And then they would refract and you'd see characteristic reflections. And you could work backwards to figure out what the structure of the crystal was. This had been applied to minerals and a lot of structure, but it hadn't been applied to proteins. When people started to look they found there were certain proteins that gave characteristic reflections. Carotin, for example. Your hair gives a characteristic reflection around 5.4 angstroms. So that suggested that there was a repeating unit somewhere in carotin that had this.

And, again, with artificial peptides sometimes they were able to see these reflections. And so that was where things stood for a while. And then one of these secondary structures, a very, very important one known as the alpha helix was deduced by Linus Pauling. Some of you have heard of him. He was a famous chemistry at Caltech. He got the Nobel prize. He also got famous later in his career because he championed the use of vitamin C to cure every ill known to mankind, including the common cold.

Although there's some merit to what Linus stated, he probably overstated some of those later findings, but his contributions to the underlying chemistry and biochemistry of proteins was amazing. And he was the one that figured out the structure that explained the 5.4 angstrom repeat.

And it was kind of an interesting story. He was in Oxford, England. And he got sick. I think it was some time in the winter. And he got bored reading detective books after a while so he thought he'd try and figure out the structure of proteins that gave rise to this characteristic repeat. So he made a simplifying assumption. He decided he'd forget all the side chains and just focus on this peptide backbone just with the peptide bond.

And he was a chemist. And he knew, what I just told you, that this had a partial double bond character so it couldn't rotate. And he reasoned that this was held together by, since these things could form hydrogen bonds that this was probably forming a hydrogen bond with a carboxyl group of some other amino acid and this was probably forming a hydrogen bond with an amino group of a different amino acid.

And so what he did was he made a sort of chain like this and he started to pleat it at the alpha carbon, which is the one that has the side chain on it, and was finding to trying to structure that would let him do this. And basically what he found was that if he made a helix that looks something like this, right-handed helix, and he could get a repeat structure that allowed him to form a hydrogen bond.

And the repeating unit was 5.4 angstroms and 3.7, excuse me, amino acids per turn. And it's a right-handed helix. It's the same sort of thing if you're trying to turn in a screw. It's got that kind of structure. And this shows you a little movie of an alpha helix. You can see this is just showing the backbone. So this is the part you can look right down the end of it.

See how you can look right though? And you can see how the hydrogen bonds are formed by turning this thing into this regular structure. And the neat thing about this then is if you put on the side chains, and you can put them on in any order, you can build a tremendous amount of diversity even within that helical structure. I think I can stop this. I just want to show you one thing, if I can manage to this when it comes around again. Stop it there.

One of the things you can see, now we're looking down the helix. And although you won't recognize the structures of all the amino acids right away, you may be able to see in this particular one. Here are a couple of aromatic rings off on this side. So this side of the helix wouldn't like to see water, and over here are a bunch of charged and polar amino acids. So you could see how you could build into a helix like that, a surface, one part that wouldn't like to interact with water and another part that would.

So that was an extremely important contribution. And there are alpha helices in almost all proteins. They'll be in little chunks coming down an amino acids chain. But they'll take up that structure. And, as you can see, it's driven by these hydrogen bonds that we've been talking about. There turns out then to be a second type of secondary structure that's important. It's called a beta sheet.

And in this case you can either line up two polypeptide chains running in the same orientation, amino to carboxyl, amino to carboxyl, or you can run them in opposite orientations, amino to carboxyl, amino to carboxyl in the other way. The latter one is called an anti-parallel beta sheet.

And if you line things up this way you'll see you can find hydrogen bonds between the chains like that. So this allows two things to form in this way and gives a sort of sheet-like structure. Whereas, that alpha helix has this tight coil like this. So over here I think we have a movie of a beta sheet. And you'll see again you can build up more than one.

Because if you look up here you can see how you are all set up to form more hydrogen bonds out in that kind of way. And, as I said, you can do this same trick putting the polypeptide chains so they have the same polarity. And so you can approximate, look at the structure of most proteins then by depicting them either as alpha helices, which you'll see in these diagrams. You've already seen a few in the examples given.

They'll look like this. Or a beta sheet which are indicated as these flat arrows. So here's a little piece of a protein made up of alpha helix, these beta sheets. What this is, actually, is a piece of the BRCA-1 gene. That's the familial susceptibility to breast cancer. The gene that causes that is called BRCA-1. And it has a special interaction domain called the BRC T domain.

This is the structure. And the only point I'm trying to make, it's of a protein that's involved in preventing you from getting breast cancer. If you get a mutation in it, or particularly in this region, for example, you can end up with an increased susceptibility to breast cancer. But what is it? It's an alpha helices beta sheet. There's green fluorescent protein. You've seen that a few times. Maybe now you'll recognize it's mostly made of beta sheets. There's a little bit of alpha helix down there, a little bit right there.

And that has the property that we've talked about of fluorescing. This is a protein I'll tell you later on that recognizes mismatches in DNA, and you get a susceptibility to cancer if it breaks. The only thing you notice here are a lot of alpha helices in it. And hopefully already your eye can begin to pick these out. This is an enzyme. What it does is it's got a catalytic ability to cleave other polypeptide chains. The functions of these don't matter. But you can see once again alpha helices beta sheets.

Here's another one. It looks just about the same, alpha helices, beta sheets, except in this case this is the human gene known as, the protein encoded by human genes called RAS. That's an oncogene. That's a gene that if it mutates in a particular way will cause the cell that has that to move a step down the pathway to cancer. So what I've done is put up a whole lot of structures that have some alpha helices, some beta sheets.

But you can get the idea that you can get very, very different biological activities from just depending on how you arrange those. OK. So there are a couple of other than forces that I need to tell you about if we're going to go all the way to understanding the 3-dimensional structure of proteins. What we can get to from that is alpha helices beta sheets.

But you saw there were loops, there were other interactions that I haven't accounted for in showing you those 3-dimensional structures. So one of them are ionic bonds. This is the third class of force. This is an extreme case of electron sharing where one atom gets all of the electrons. So aspartate, which I had up on the board, aspartic acid looks like that, but under physiological conditions the oxygen will get all the electrons and you'll have a hydrogen on it.

And a consequence of that then is that if you have a polypeptide chain that over here has an aspartate and over here has a lysine, which is the four methylene groups, and the positively charged thing here, you can get an ionic bond between those two amino acids that can be very far apart on the polypeptide chain.

There may be a lot of amino acids in between, but what they then do is bring these two points together and hold them like that. The next class of force is kind of tricky. You may have heard of it in chemistry. It's referred to as van der Waals interaction.

And the basis of this, without going into it too much, is even a nonpolar bond --

-- can have a transient polarity. And this then induces --

-- a transient polarity in a nearby bond. And it has to be a really nearby bond. So about 0.2 to 0.4 nanometers. Remember, covalent bonds are roughly half that distance or something. So it's got to be a very, very close interaction. It's weak. It's only one-third to one-quarter of a hydrogen bond, which you may recall is about one-twentieth of a covalent bond.

But there can be many, many of them if the surfaces fit together really, really tightly. So if you have a protein fold, so there's a surface here, and then it folds up in such a way that there's a surface here, then you can get a lot of van der Waals interactions down here. Now, I've never had a really good way of explaining this. But today, part of these activities of this Hughes Professorship, I've set up some seminars on teaching.

And I've invited a guy from Berkley named Robert Full who is talking in 68.180 at 4:00 PM. And I borrowed some things from him this morning. And we're just going to take a quick tour because I want to show you this. He works on, well, he does a lot of things. He works on biomotion and how animals work.

But one of the things he works on, let's see if we can get this guy to go here. Oops. How do I figure out how to get it to play here? Hang on a second. I just discovered that the PowerPoint is not really terribly effective.

So this isn't working as nicely as I would like. OK. Let's try this. Just a minute. Where are we? Here we go. OK. Let's see if I can get this to go. So he studies a bunch of things, but he did an undergrad project library studying geckos. And here this is a transparent surface. And he's studying how the geckos climb up and down the thing.

And they were making measurements. And they found they couldn't account for why this was such an efficient organism. It used much less energy than most things, so they started looking into how it adhered to the surface. It can go up a vertical wall, as you can see here. And so they were able to look underneath and they could see, see how it sort of peels off the surface? And this was a robot that they eventually built that's not using the same molecular bases but uses this peeling thing. And they can get a robot that climbs up a wall.

But that's not what we're going to talk about here. We're going to instead, I hope, go back to here. And you can see that all of the geckos have these sort of bizarre toes, and so they started looking to see what the underlying principle of this was. And they saw it has these setae. And they got looking in greater detail and blew it up.

And then they found that there were, as they started looking there were these little hairs. And that's a 900 fold magnification. And once they got looking in more detail they found the ends were split so that they were, the very ends are about 200 nanometers roughly at the end of this. And so a gecko has about a billion of these on its feet. And what it turns out it does --

And just to see, here's a human hair. You see how it splits down? Now, this is made of keratin, the molecule I just mentioned that was used, alpha helices, but it's very, very fine. And what it can do, it can make van der Waals interactions. This is an animal that sticks to the wall by van der Waals interactions. And the peeling away allows it to break those bonds. But, as you can see, they're enormously important.

He's got here a micrograph. They're measuring the force, and the force is just huge. This is the end of the thing, the frayed end sticking to a surface. And for those of you who didn't think biology any relevance to you, Bob was telling me about this. They followed up, he's an engineer as well and builds interdisciplinary teams, and they've measured this stuff. But this is turning into what appears to look like it's going to be a \$30 to \$50 billion industry as all sorts of things are --

They're beginning to realize it can hold car parts together, it can go in space shuttles, Post-it notes. And here's a little Band-Aid they made. They own the patent on this self-cleaning dry adhesive. It doesn't have to be made out of gecko stuff. It could be made out of all sorts of things. But, anyway, here's an example of where not only are van der Waals forces very important, but where somebody who started a very simply aspect of biology worrying about the efficiency of how geckos ran and pushed it all the way down to the molecular level understood a principal that's going to make somebody a very large amount of money.

OK. The last, and if anybody wants to come, he's an amazing speaker. Perhaps one of the most exciting speakers I've ever heard. 68-180, 4:00 PM if you want to go. He'll have more of that sort of stuff to show you then. OK.

So the last force here, it's not really a force, but what we'll call hydrophobic effects. And what I mean by this is that the principle of this is that amino acids that don't like to interact with water, so --

So hydrophobic amino acids. These are ones like leucine and phenylalanine. Well, I showed you the water the other day and how it was forming hydrogen

bonds between the molecules. So if you're going to stick another molecule in there, you're going to break a bunch of bonds. And if you're not charged or polar you cannot make new bonds with the water. And so what happens, if you put these together, just like if you put oil together it will all bundle up and it will minimize its interactions with water.

And that's what proteins do. Here's the structure of a protein all folded up in 3-dimensional space. And you can see at the core of the protein how there are these many hydrophobic amino acids that are interacting. And let me just, I'm going to close by showing you one more little movie. And the new version of PowerPoint doesn't do this well so I'm just going to get out of this for a second here.

This is a really cool movie I saw. I want to show you a DNA repair protein sticking to a piece of helix. Can you hit the lights somebody there? So this is a lesion on a piece of, see the double helix here?

And what I especially liked about this is this is sort of a Star Wars movie. You're going to fly down the major groove of a double helix. And you can see where this particular protein folded up in 3-dimensional space is reaching down into that helix. So this is sort of putting together the two things that I've been telling you about. This blue is a DNA repair protein. Oopsy daisy.

A DNA repair protein that's able to find a lesion in the DNA. And here's the double helix that's the two chains held together by hydrogen bonds. And then, as you can see, there's a groove on each side. And the protein is searching down into that groove.