>> So it is again, March 2nd this time it's 2015. And tonight, continuing this story entitled Cellular Anabolism, specifically the production of protein. Hard to overestimate its importance, because what makes one cell different from another is the kinds of proteins that it can or cannot make. And so we say proteins determine the physical and functional characteristics of each of your cells. Proteins are not made haphazardly. They're not made spontaneously. They're actually orchestrated, that is their production is controlled by the nucleic acids namely DNA and its subservient partner RNA. We said that DNA contains all the information for the synthesis of every conceivable protein that the human body can make. And these instructions are defined in or part of what's called gene. Everybody's heard that term and fewer people understand exactly what it means. A gene is that portion of DNA with the information to make what? One specific protein. And so we mentioned that maybe on the human genome, yours, mine, there might be 20,000 actual genes. Nobody knows the exact number even today. So DNA works with his partner, RNA, in ways that we'll discuss tonight. But it's RNA that does the dirty work. RNA does the heavy lifting. RNA actually brings and sequences amino acids out here in the cytoplasm to produce a final protein. But if that protein is faulty, whose fault is it? Is it the RNAs or is it the DNA in most cases? Usually the DNA, because if you have faulty DNA you're going to have faulty RNA and therefore a very different protein, which can be catastrophic in terms of health or even survival. So there is where we were as of last Wednesday. Let's move forward now and gently and slowly with introduction to this masterful molecule, this iconic symbol of life namely DNA, deoxyribonucleic acid. You can't turn around without seeing it on TV or magazine covers. It's just so engrained in our culture now a very beautiful molecule actually. And you might not recognize these chaps. This is the picture taken in the early 1950s, but these two gentlemen are credited not with discovering DNA. They didn't discover it. It had been known for some time. What they did is determine its shape, its structure, the classic double helix. This is Francis Crick. This is James Watson. He was a mere baby at a time and he still alive today. Francis Crick died about five or six years ago. So are these guys famous? Did they win a Nobel Prize? Should they be revered? Absolutely. This is the sketch that Francis actually drew on a napkin, which ultimately led to the verification of DNA as a double helix. So, quite an interesting story if you ever want some science intrigue, this is Francis Crick, James Watson. And they did win the Nobel Prize in conjunction with Maurice Wilkins in 1863. An interesting back story is this of Rosalind Franklin. She was a, well, quite a groundbreaking pioneer because were they many women scientists of her caliber at that time. But actually it was her photograph of DNA, her x-ray diffraction photograph seen here, which actually provided a powerful clue for Watson and Crick. And you say, well, "Then, why didn't she share the Nobel Prize too?" Well, sadly she died in 1958 of leukemia, interestingly. Probably due to her exposure to the very x-rays that she was using for this photography, this photography. But anyway, because she had died, the Nobel committee had a policy against awarding any Nobel Prize to someone who has passed away, so quite an unfortunate story but I wanted to give her due credit. So here's a more recent picture of Watson and Crick. Actually it's not that recent but at least when they're both were alive and speaking to each other. So, there you go some little biological history. And let's move forward, DNA. It is an acronym for deoxyribonucleic acid. And as such, it's a very long polymer of subunits known as nucleotides. What's a nucleotide? A nucleotide in this case is a deoxyribose sugar. A PI, what's that stands for?

[Inaudible Remarks]

Inorganic phosphate and one of four possible so-called nitrogenous bases. Now this illustration is by no means DNA but it does show the basic structure of a nucleotide. This then is the sugar. The sugar in this case is deoxyribose. These are the phosphate groups which link these sugars together as you can see. And the organic or the nitrogenous basis, four of which, some of which are double-ring structures called purines. Others are single-ring structures called pyrimidines. So, it is in fact not a single or linear molecular but a side by side double helix. In other words, a ladder which is twisted as this model here is. So I pulled apart one of the subunits. And tomorrow in lab, indeed, you're going to assemble some DNA using models like this on a smaller scale. So if this is a nucleotide what are the parts? What are the three parts of any nucleotide? And let's say the sugar is this black piece and therefore this white piece would be the linking unit here known as a phosphate group. And then in this case the colored piece would be one of four possible bases. So the double helix, a ladder which is twisted, and here we see a sort of three-dimensional attempt to show that relationship. SP-SP-SP, what's that? S?

>> Sugar.

>> What's the sugar? The sugar is deoxyribose. P stands for? Phosphate group. And what makes the molecule unique

and important as an information source, is not the SP-SP which is boring and redundant, but rather the sequence of basis. Notice that these bases are held together. Here we flattened the molecule down. So we can see that in every case, there's a mandatory relationship between pyrimidines and purines. So cytosine will always and only pair with guanine. And thymine will always and only pair with adenine. And when we say pair, we mean form a bond. What kind of bond? What kind of bond? H bond. Yeah, A weak hydrogen bond. So from where you're sitting you can see this white piece here which is linking these bases. So these are the hydrogen bonds. Weak enough to allow for separation, and that separation is key in the ability of this molecule to work as a template as a source of information. Interesting thing about DNA, the unique feature of DNA, is that it's self replicating which means simply it can make more of itself. There aren't any other molecules that can make more of themselves. And indeed it does this as you'd imagine in the course of cell division. And indeed, tomorrow we're going to be talking about that very process called DNA replication. So deoxyribonucleic acid, not discovered by Watson and Crick but the structure was illustrated by them. An interesting remark that James Watson was reported to say, when they actually built their model, which was not too different from this. He said I knew it would be a beautiful molecule. And indeed, it's hard to imagine a molecule being beautiful maybe, but indeed I think it kind of is. What's the sister molecule related to DNA? Ribonucleic acid. Essentially subservient to DNA, indeed made from DNA, not the entire molecule but a piece of it known as gene. But still it's NA, what's NA?

>> Nucleic acid.

>> Nuclidic acid. This molecule is not made of deoxyribose but rather ribosugar. And it's not a double helix in most locations. That is it tends to be more of a single helix. Imagine a slinky and just stretching it out, that's a single helix. But as before, this is the sugar. What is the name of this sugar this time?

>> Ribose.

>> Ribose. And these are the linkages, the phosphate linkages. So as you read from top to bottom, this would be phosphate sugar, phosphate sugar, phosphate sugar, on and on and on. These are the nitrogenous bases. Why are they called nitrogenous, because they have a lot of nitrogen. They're also called organic bases, and they come in four possible forms, again, pyrimidines and purines. But an interesting substitution occurs here. Yes, we have cytosine and yes we have guanine, but in RNA thymine is replaced by a molecule which has the same basic performance, that is substitutes for, takes the place of thymine, it's called uracil, but uracil is found obviously then only in RNA. We show these lines suggesting there's a pairing potential here that is pyrimidines pairing with purines. But as a remark, we've already said that RNA does not form in most locations, a double helix but rather exists as single helix with no regular base pairing. However, the molecule at times may be so long that it actually loops back, folds back over itself and therefore the end of the molecule may form hydrogen bonds with other earlier portions of the molecule. That might seem complicated but it creates a kind of looping appearance, at least to the so-called tRNA molecule. Going through this task because I bet you heard it before, you've all had biology and therefore these notions are not entirely brand new. And if they're not, what are the three acronyms, the three functional types of RNA? M stands for?

>> Messenger.

>> Messenger, method or function that will give some attention to. TRNA, the T stands for transfer. And R, rRNA is ribosomal RNA, an RNA that actually helps in the construction of these organelles we'll see tonight called ribosomes. This illustration here looks a lot like DNA at first glance, but if you let your eyes do some tracing, you'll see this is a single helix, this just folded back over itself and so this happens to be representation of tRNA, which has some areas of base pairing based upon this looping back on itself. So before going on, let's just sort of sit back and compare and contrast these two molecules. Which of these two nucleic acids is limited to, found only in the nucleus? DNA. RNA is created, that is built in the nucleus, but does all of its actual function out in the cytoplasm. Which of these molecules contains all the information for any protein that a cell might make?

>> DNA.

>> DNA. Which of those molecules carries information for the synthesis of just one protein? RNA. So RNA, again, essentially a copy of a gene in ways that we'll see. What are some other distinctions? How is RNA different from DNA?

>> The uracil.

>> Yes. So uracil takes a place of thymine. Good, what else?

>> Absence of oxygen.

>> I'm sorry?

>> Absence of oxygen.

>> Well, yes. The sugar in RNA is what? It's ribose, not deoxyribose. Good. And what about the tertiary structure, the overall--

>> Single helix.

>> Right, single helix versus double helix. So those are important facts because they relate to how these molecules work, that is work together for protein synthesis. Now, certainly when Watson and Crick determined the structure of DNA, the real breakthrough was determining how this molecule could contain and code for the information to make proteins. It was instantly clear to them that the so-called information could not be represented in the sugar or phosphate, because that's redundant and boring. It is after all phosphate sugar, phosphate sugar, and so on. S, it became clear that the code had to be represented in the sequence of bases, along one or maybe both strands, they didn't know at the time. And so from their discussions and early speculations, emerged what we now accept as the universal genetic cod, universal because it applies to you, me, your cat, a tree, anything living on this planet. And it turned out as predicted, that is the sequence of the purine and pyrimidine bases which really ultimately determines the order of amino acids that will be assembled in a given protein. But there in lied an immediate dilemma. How many different, how many different bases are there, let's say in DNA? And what are they? Adenine, guanine, cytosine, and thymine. How many different bases? Just four. How many different amino acids are known to be necessary for and often included in any given protein?

>> Fifty.

>> Well, 50 is a number you recall but 50 referred to the minimum number of amino acids in even a small protein, but the number of amino acids to choose from, are 20. So, why is that a quandary? Why is that a problem? How many bases are there?

>> Four.

>> How many amino acids? So how could an alphabet made up of four letters code for 20, 20 different amino acids? So, why would a genetic code that is read, that is involving bases that are read as individuals or even grouped into pairs be unworkable? Let's look at that real close. Why would a genetic code reading basis individually be unworkable? How many bases are they?

>> Four.

>> How many amino acids are there?

>> One.

>> There's your problem. We have these bases. We have these amino acids. Clearly, a single base could only code for one of these amino acids. So, four could only actually select four or code for four possible amino acids. So I'm sure it as Watson and Crick who thought, well, maybe they're not read individually, maybe these bases are grouped as pairs. And why is that unworkable? How many different ways can we group four bases in groups of two? We can group A with A, A with C, A with G, A with T, T with A, T with C, and so on. And how many combinations can you actually come up

with?

>> Sixteen.

>> Sixteen. Why is that unworkable? Still 20 different amino acids. So apparently, these bases are not read individually, they're not even grouped or read as pairs because both ways would not cover, that is not allowed for coding for all different amino acids. So, with that said, their instant mathematical minds must have said, "Well, OK, these bases are not read singularly, they're not grouped in pairs, they're must be grouped in, well, triplets as it were." And it turns out, how many combinations can you get then? How many different ways can you group these four bases in triplets? Well, 64. And is that more than enough to cover the 20 known different amino acids? So, they presumed and they were quite right that this is the way it works, bases are not read individually. They're not grouped in pairs. They're grouped and interpreted in three letter words, called triplets, each coding for one amino acid. But wait, how many amino acids are there?

>> Seven.

>> How many different combinations are there? So that's a bit of a quandary because, well, that implies that maybe some of these triplets, code for the same amino acid, and it's maybe possible that some code for no amino acid, certainly in the early years of their investigation of this, it wasn't known. But today we can list all 64 DNA triplets, and we can actually tell you with confidence the amino acid that they code for. And this is only a partial list, and nothing you want to memorize. But it's up here for a reason. These are names of amino acids. How many different amino acids are there?

>> Twenty.

>> And notice over here on the left, we have, what is this? This is triplet which reads from left to right, GGT. What amino acid is that ultimately code for? Proline. But also what? GGG codes for the same amino acid. So are some of these triplets redundant, that is do they code apparently for the same amino acid? Yes, perhaps unexpected but possible, and indeed the case. So, can one amino acid be represented by more than one triplet? Can one amino acid be represented by GGT and also be represented by, in this case, GGG? Yes. Why? Well, just because, because there are how many possibilities? Sixty-four. And so that allows for this possibility which turns out to be true. And then, can one triplet, though, represent more than one amino acid? This question might sound like it's the same question. But we're asking this, can GGG code for proline and might it also code for something else? The answer is, never, no. And that would be catastrophic. In other words, do you want the word DOG to mean dog today and elephant tomorrow? No, you want there to be consistency. So, what about this statement? Can any given triplet represent more than one amino acid? Never. And the reason that it can't be and shouldn't be is that it would create chaos. That is a code would mean something today and something else tomorrow, would be constantly morphing into something else, be probably pretty ugly. So that's interesting fact. Shelby [assumed spelling]?

>> Can you say the reason why--yeah, that.

>> Well, why? Are there lots of possibilities here and only 20 amino acids? So, it's not like this is purposeful, it's because the math allows it. How many possibilities here?

>> Sixty-four.

>> How many different amino acids? So is it possible, indeed true, that some of these triplets indeed code for the same amino acid? Don't try to extract a reason for that. There is no real reason other than it's possible and it turns out to be true. It doesn't accomplish any purpose but it is an interesting fact. Another interesting fact is that, some of these triples code for nothing, and that's also a possibility. That is they simply don't code, and it turns out that they serve as punctuation. That is, what are called stop messages, which basically initiate either the start or the stop of a particular protein's construction. So this is indeed universal. And by that, I mean, it's the same for you and me and the elephants and anything on this planet, an amazing fact in itself. Notice we've left out some information here, information you could fill in because these triplets as we'll soon see serve as template for codons and anticodons, concepts that we haven't even discussed or defined at this point, but perhaps you're already aware. You go back then and fill this

information in when we get further along in the start. So, let's go with this story, which is of course protein synthesis. Let's just pause and emphasize the case for protein synthesis. What functions do proteins serve that make their synthesis so valuable? Our enzymes protein, or some hormones proteins, are there important proteins in the structure of cell membranes, are there neurotransmitters that are proteins? And the list goes on and on. And so, protein synthesis can't be overestimated or overemphasized. From biology, you know then that protein synthesis is a two-stage process. Stage one occurs in the nucleus, stage two takes place outside the nucleus in the general cytoplasm. The name of what goes on in the nucleus, it's called what?

>> Transcription.

>> Transcription. A word that makes sense, because what's a transcriber? Like the profession. They are people who are transcribers and don't they change one form into another, maybe they take the spoken word and put it into the written word. But the idea here is simple, we're taking DNA using it as a model and we're going to copy not the entire the molecule but just that portion associated with the synthesis of one protein. In other words, we're going to copy a gene, not as a gene but as a molecule which will essentially allow for protein synthesis, something called RNA. So I animated this and certainly YouTube and other animators do a better job, but we'll do what we can to show the basic steps. What's this molecule shown here in a flat two-dimensional image?

>> DNA.

>> DNA. So, not shown along these blue ribbons, not shown are SP, SP, SP, SP, what's that?

>> [Simultaneous] Sugar formula.

>> And the sugar again?

>> [Simultaneous] Deoxyribose.

>> Deoxyribose. The sugars and the phosphates are unimportant, what gives this molecule any information that is what codes the information is the sequence of bases. These bases are paired in the manner that we said. And we know now that there are three billion, three billion, with a B, three billion base pairs along your DNA molecule. So obviously, they didn't have room for three billion, so I shortened it out for the sake of discussion. And let's do this. Let's show as best as we can, transcription, which begins with the identification of and the separation of the portion of the DNA with the information we're after namely a gene. And incidentally, even a small gene is usually 20 to 30,000 bases long. But OK. Step one, the molecule has to unwind, it has to unzip not from one end to another but only along that segment which contains the so-called gene. And so here we've done it in sort of primitive animation. We'll show it again. There it is, there it is. So it's unraveled at least along a certain length of this molecule. And by unraveling, what do we mean? What bonds are broken to separate these previously connected strands?

>> Hydrogen bonds.

>> Hydrogen bonds. And now we show and reveal SP-SP, just for clarity. But these colored segments, these colored symbols here are the bases. And remember, they are adenine, guanine, cytosine, and thymine. So here we see represented without letters, these pyrimidines and purines. It turns out that only one of these strands will be involved in this process, and that one is called the template strand, let's assume it's here. And the template strand will, indeed, be a template that is used to make this RNA molecule which will ultimately leave the nucleus. So we filled in some letters, G, T, A, C, C, and so forth. Step two, involves the recruitment and the alignment of free, what free?

>> [Simultaneous] Ribose--

>> Ribose nucleotides, not deoxyribose but ribose. Keep in mind, what we're making is not more DNA, we're making RNA. And so, the subunits are ribose nucleotides. And they will align, that means they'll position themselves spontaneously. That is they will link up according to hydrogen bonding. And that will occur only on the template strand not the other strand shown. So--Oh, that was cool. Let's do that again. So, what is this thing which just flew in from

nowhere?

>> It must be a nucleotide.

>> It must be a nucleotide. And what is the sugar that's not identified here? Ribosugar. P stands for phosphate. And this guy must be--well, must a base. And we can predict what it is because if this is guanine, this would have to be--guanine always and only pairs with cytosine. Oh, it bring this in. All right. So, these are what? These are free ribose nucleotides. Are they connected yet to each other? Is this nucleotide bonded to the next to it at this time? No. What's holding these nucleotides in place is nothing more than the invisible weak electrostatic force here, which we've already said is a hydrogen bond. In this case, between cytosine and guanine, in this case, adenine and thymine, what's this?

>> Uracil.

>> Uracil, yes, because thymine is absent in the nucleotides used for RNA. So, everybody good with step two? Alignment of free ribose nucleotides only along this so-called template strand, and in accordance with the rules of mandatory base pairing. Oh, great. Let's finish it off. Next, apparently three. There's got to be bonds form now between these independent nucleotides. And how or between what are these bonds created? Are we going to bond sugar to sugar, base to sugar, or phosphate to sugar?

>> Phosphate to sugar.

>> Phosphate to sugar. Phosphate to sugar, phosphate to sugar, phosphate to sugar. And those are indeed bonds, covalent bonds. And the creation of these bonds is not without the utilization of energy. What is the energy currency that would be used of expended to create and therefore establish these bonds?

>> ATP.

>> ATP. So this step is going to gobble up a lot of ATP. This process that we're about animate is called nucleotide polymerization. To polymerize means to hook a lot of similar molecules together, in this case, the RNA nucleotides. All right, you probably blink so I'll start over. What are we making here? Bonds between what?

>> Phosphate.

>> Phosphate groups and ribosugar. Watch it. Stunning animation. This is breathtaking. The power of powerful and--oh, OK. Anyway, those are sugar phosphate bonds. Now remember, the bonds that are not shown here are still in force. What were--these bonds linking these bases together and they're relatively weak as you know, therefore this molecule can literally peel away from the DNA molecule, and will do so. But before we go on, this step, number 3, nucleotide polymerization, uses a great deal of ATP to create these sugar phosphate bonds. And it also requires an enzyme which is not shown here but definitely involved that needed, its name is RNA polymerase II. That's an 11, its RNA polymerase II, Roman numeral II. Now, what RNA polymerase II does, is create these bonds that I've shown in red. Not by itself. Remember, enzymes don't do that but certainly we had used some ATP. The important thing about this enzyme which is eluded to here is that it's DNA dependent. What does that seem to mean? This enzyme won't work without the presence of?

>> DNA.

>> DNA. And that should be instantly appreciated, because would you want to take nucleotides and just hook them together any way without any respect to or guidance from DNA? Would that make RNA? Yeah, but that'd be like just making alphabet soup and then just pulling out letters randomly and hoping that it a word beheld. So, why do we care, why are we glad that this enzyme is DNA dependent? It's only going to work if what's here? That guarantees that what you're making is in fact, is in fact a copy of something meaningful, that is a copy of the gene. So it couldn't be more important that this enzyme is DNA dependent. At this point, the molecule which we've represented only in a short piece here is called the primary RNA transcript. And literally it peels off the DNA. And incidentally, what supposed happens to the DNA after this parting of company occurs? Has the DNA been altered or mangled or changed in any way? So, is

this RNA going to reform the double helix in this location? Sure. So the DNA is not consumed or destroyed. And what we've made is a copy, not of the entire molecule but just of the segment. And that segment at least at this point is called the primary RNA transcript. Probably, easily, hundreds of thousands of basis long. For the longest time, it was assumed that this primary RNA transcript was an accurate and complete copy of the gene, which in fact it is. And it assumed that this would then move unaltered out into the cytoplasm and begin and engage in the next process which we haven't discussed yet, but the final process called translation. But it turns out that that was, well, false. That is it turns out that this primary RNA transcript contains a lot of sequences which are basically garbage that is nonsense segments of the DNA. And so a very important next step, number 5, has to occur and does occur before this molecule leaves the nucleus. And it's called RNA what?

>> Splicing.

>> Splicing. Now, this analogy goes back to the days of audio tape which nobody remembers. But years ago, there was recording tape that was unreal. You might even recall those little square boxes that had tape in it. Remember those--some of you do. They were call cassettes. All right. Anyway, RNA splicing borrowed that concept because in their recording industry of the film industry, when you are editing the movie, you basically take a pair of scissors, don't you? And you do what, literally? Cut out the stuff you don't want and then put together the images or information that you do want. So, literally, that's what's happening here. The primary transcript has to be proofread and essentially introns, a word standing for noncoding segments have to be identified just like you would proofread something that you wrote or typed up on a computer. And those words which were misspelled or inadvertently included are going to be literally cut out with enzymes. And then the useful portions, the meaningful portions, the coding segments are then going to be enzymatically spliced back together. So in fact, this proofreads and reduces the size of the RNA transcript considerably but basically gets it ready for its final ability to be used in translation. So step 6, the processed or edited messenger RNA, will now diffuse out into the cytoplasm and pick up our final process there. We talked about RNA polymerase. RNA polymerase, what? Two, which implies that there's probably a one and maybe a three. And so, indeed, remember, messenger RNA is made under the influence of this enzyme RNA polymerase II but transfer RNA is also dependent on an enzyme. One called RNA polymerase III and ribosomal RNA is also coded that is created by a different enzyme RNA polymerase I, each of these then producing a different kind of RNA. All right. So let's just summarize before we run on. DNA. We're going to--what's the word, transcribe it. Do we need to copy it all? Nope. Just that portion that has the information for a protein, that portion is called a gene. The gene has to be identified and the helix in that area has to be unwound. What bonds have to break to allow that to occur?

[Inaudible Remark]

And then are both strands copied or just one? The one that's used is called the--

>> Template strand.

>> Template strand. What raw materials do we bring in to essential form our developing RNA?

[Inaudible Remark]

What kind of nucleotides? Ribose nucleotides. How many different kinds are there? Four. Some are uracil carrying, some are adenine, some are guanine, some are cytosine. And they come and align against the template strand. What dictates their position or location along that strand? Uracil always pairs with?

>> Adenine.

>> Adenine. Guanine always pairs with cytosine and so forth. What enzyme is used to link these nucleotides together to form the sugar phosphate bonds?

>> RNA polymerase II.

>> RNA polymerase II. Is enzyme necessary? Yes. This enzyme is DNA dependent, which means it won't work unless

what's also in the vicinity?

>> The DNA.

>> Guaranteeing that what we're making has a meaning, some accuracy and therefore be useful. ATP used in process?

>> Yes.

>> You bet, by the bucketful. And what we end up with is a primary RNA transcript. It has to be pruned. Do you what I mean? We use the word edited or spliced. And that means we have to identify--not we but enzymes look for and identify the noncoding segments. Noncoding segments are called introns. Those are cut out, thrown away, and then the exons then are connected enzymatically to create a meaningful, a meaningful RNA. What kind of RNA have we made? Well, actually, we've made rRNA and tRNA previously but what we showed was the creation of M, mRNA which is messenger RNA. So now, it moves out to the cytoplasm. Great. And what's going to happen there is it'll partner up with existing tRNA, and these organelles called ribosome which, themselves, are made of RNA. And basically we're going to gather up amino acids. How many different kinds are there?

>> Twenty.

>> Twenty. And we're going to hook them up in the order specified by the RNA. Remember, this code is read and grouped not as individual basis, not as pairs but as triplets. So every three basis will symbolize or represent a single amino acid. So, unfolding the story in chronological order, mRNA, the completed and pruned mRNA will move from the nucleus to the cytoplasm. And here it is laid down in an arbitrary partial illustration. SP, SP, SP, what's that mean?

>> Sugar.

>> And these are obviously the bases. How are these bases grouped in red individually?

>> No.

>> Pairs?

>> No.

>> Nope. Triplets. And so, the first triplet in this example apparently is--reading from left to right, A, U, G. The next one is U, G, U, and so on and so on. So, next, we have to involve these organelles made themselves of RNA and protein. The organelles are called ribosomes.

[Noise]

I threw in the sound just to wake up some people who were snoozing. What was that?

[Laughter]

Aliens, OK. Now, I don't want to play it back, it was too funky. But, all right, this is what? The ribosome. Remember, the ribosome is not just a blob or this shadow but rather an organelle. An organelle made of RNA and protein. From biology and textbooks, you've probably went across the analogy that ribosomes are the workbench, because this is where the work of translation is going to take place. And what about that very word translation? What does that mean in any ordinary conversation? I'm going to say please translate that. Meaning, convert it from one language to what? Another. And so, the language that we are beginning with is a language of base sequence and we're going to change that into a sequence of amino acids. So, we're translating from the base language into the amino acid language. Attachment then of the mRNA to the ribosomes. And remember, these bases are not read into individually, not in pairs, but in triplets. And there's a name for triplet found on the mRNA, and that name you recall from biology is called a codon. So, with this example that we're going to use, what is apparently the codon, the first codon right here?

>> AUG.

>> AUG. The next one is UGU and so forth. So, as you would know, there are 64 different codons just as there are 64 different triplets. Some code for what?

[Inaudible Remark]

Some code for no amino acid. And this then must be apparently the so-called start codon AUG. Start meaning it represents the beginning of this process, in this case called translation. Now, what's the other molecule, the other form of RNA that's mentioned here but not yet shown? TRNA. And tRNA means a molecule which is basically a taxi cab that is transporting, as it turns out, amino acids from here to this particular location in the ribosome. So here comes a tRNA with a three base sequence showing here as UAC. Why would we expect this to attach at the location? What's U stand for?

>> Uracil.

>> And uracil always bonds and pairs with adenine. Adenine always bonds and pairs with uracil and C always with G. What is the attraction? What's holding these together? It's hydrogen bonding. Notice that this tRNA molecule does not carry anything and therefore is essentially initiating this process. But transfer RNAs normally do transfer amino acids. That is their whole job is to pick up specific amino acids. So, this little representation is meant to represent a tRNA and at one end the molecule or at least one location in the molecule, there are three bases which are unpaired and served to match or otherwise align with the codon. And for that reason, these are called the anticodons. At the other end of the given tRNA, there's a receptor, an enzyme which is in this case, identified here as aminoacyl tRNA synthetase for a particular amino acid, in this case cysteine. So what's going to attach there is CYS which is an acronym for cysteine. What made that happen was the specific enzyme which would only allow cysteine to attach there and therefore only be carried away with this particular tRNA, which has this apparent sequence ACA. So, as you would expect then, there are 20 different enzymes, 20 different specific tRNA synthetase enzyme used for the attachment of the 20 total different amino acids. What we end with is this molecule here minus the enzyme, which is given the name what? Aminoacyl tRNA complex. In other words, the tRNA attach to a specific amino acid. And this portion of the complex is important because that three-letter base sequence will serve as an anticodon which will attach at selective locations along the mRNA sequence. So watch this because it's stunning. OK, you might have missed it. Where is this guy going to go? What is this anticodon as we read from left to right? ACA. What is the codon, the only codon that would form a hydrogen bond with those three bases? U--What?

>> G.

>> G.

>> U.

>> U. Oh, and that happens to be by no accident, the next codon. So, there you go. What have we done? We've managed to position to bring in to this process, what amino acid? Cystine. How do they get there? That is we brought it there was this tRNA molecule. And incidentally, the attachment of that amino acid--bless you--was not without energy use. That is ATP was used to make that happen, and so this process as we'll soon see is also very ATP dependent.

>> I have a question.

>> Good.

>> So the one next to tRNA [inaudible], anticodon because it doesn't have the amino acid?

>> This is an anticodon, true, but this anticodon doesn't code for any amino acids, so it doesn't really bring anything in but it does serve to initiate to start the sequence as the AUG does too.

>> So the anticodon is like specific [inaudible]?

>> Well, remember, there are how many different codons? How many different codons? Sixty-four. How many different amino acids?

>> Twenty.

>> So do they all code for amino acid? No. Do some code for the same amino acid?

>> Yes.

>> Yes. This one happens to code for no amino acid and therefore essentially begins this process. Hang on, we'll finish this up. What's the next codon? The next codon is GUC. This is quickly going to get boring, because once you get it, you know, it just then automatic. What is the only anticodon that could and would form hydrogen bonds?

>> [Simultaneous] AUG.

>> All right. And so as we've already said, the three-base sequence on the tRNA is called the anticodon, which will form hydrogen bonds only with a complimentary three-base sequence on the mRNA, which of course contains the codons. So, in short, where are the codons?

>> [Simultaneous] MRNA.

>> MRNA. What carries the anticodon?

>> [Simultaneous] TRNA.

>> TRNA. And tRNA's job is not just to form these bonds, but rather bring in an amino acid and therefore position it along this developing amino acid chain. So, in green is the codon, in yellow is the anticodon in this story. Oh, that's interesting. What happened there? This start molecule here, this start codon, in this particular tRNA basically what? It flew away. OK. Well, that all right. And what's this big shot over here called again?

>> [Simultaneous] Ribosome.

>> Clearly, if this is where these events take place, there's got to be a physical movement. This got to chug down this track or if you wish, the mRNA is going to pass through it, regardless of where you look at it there's going to be some movement there. The next, the next, what is it? The next codon is GUC in this case. And the only anticodon would be CAG. This brings in a totally different amino acid, this happens to be valine. And OK, so what? Keep in mind that our goal, we're trying to make protein, what's protein made of?

>> [Simultaneous] Amino acid.

>> And we have how many so far?

>> [Simultaneous] Twenty.

>> And they're not even linked yet, right? So, so far we've got two amino acids, shoulder to shoulder, but nothing more than that. Clearly, there has to be a bond created between these two. And you already know the name of a bond between amino acids in any and all proteins. Those bonds are called? Peptide bonds, special kind of covalent bond. It turns out that the ribosome contains or is actually in part, an enzyme called peptidyl transferase which catalyzes, that means promotes the formation of a peptide bond between these now adjacent amino acids. So, what are the two amino acids we've got lined up so far? Cystine and valine. And now we're going to create that bond. In case you blinked, OK, we'll show it again. Oh! And at the same time, what happened apparently? Oh, in case you missed it. What this molecule

here?

>> [Simultaneous] TRNA.

>> That tRNA is going to disappear. And the next time we see that tRNA, if we see it again, it would be bringing in another cystine somewhere else along these lines. So simply put are these tRNA molecules recycled? That is they go out and pick up these amino acids endlessly. What a boring career. But OK, they're just molecules. What we got next? Well, the next, the next codon seems to be AAG. And notice the ribosome moved in that case. And we would expect what to come in there?

>> UUC.

>> UUC. And it's bringing in a totally different amino acid, abbreviated PHE which stands for phenylalanine. You don't need now this. But clearly it's a different amino acid, yellow. And what are going to do next? I mean come on, it's redundant now. We got to create a bond there. What do we call that bond?

>> [Simultaneous] Enzyme.

>> And what enzyme makes that happen? And that enzyme is actually part and parcel of this organelle called the ribosome. Does this use energy as suppose?

>> Yes.

>> Yes, all right. So once again, through the magic of my fingertips, all right. And how do the amino acids have we linked so far? Is this a protein yet? No, it's just three stupid amino acids linked together. After all, even the smallest, most pathetic of proteins is 50, right, so we got long ways to go. Don't worry. We're not going hold you for every one of those 50, because you already got the idea. But let's finish it out at least as far as we can go. We've already revealed that as these peptide bonds are made, what happens to the tRNA molecule which delivered that amino acid? All right. So it's freed, that means it detaches from its connection through hydrogen bonding and it's going to disappear and pick up that same amino acid for some other attachment later on in this sequence. So, all right, what's the next one? It seems to be what? I'm looking at it CCG, and the appropriate anticodon would be GGC. In this case, bringing in something called pro which is proline, a different amino acid. And as before, what's going to happen next? All right, the peptidyl transferase is going to catalyze that bond, and the tRNA disappears and now we have, how many amino acids together?

>> Four.

>> Four, OK. Next one. UUG which would only attract and make a complementary linkage with the anticodon AAC, that's bringing in Leu which stands for leucine, another bond, OK. Next, notice the ribosome moves on, next one seems to be UGG. That would mean ACC bringing in tryptophan in this particular case, bond formed, tRNA disappears. Oh, now because it's getting boring, we cut to the chase here. This final codon as you can see reads UGA, which codes for a so-called stop, it is in fact to stop codon and the tRNA which attaches their carries no amino acid, symbolizing the end of this process which in life would not be this brief. Remember, even the smallest protein contains what?

>> Fifty.

>> Fifty. And we've only got a handful here, just for simplicity sake. So, that final tRNA carries no amino acid, and essentially symbolizes the end of this process, previously identified as translation. And now, these amino acids are free to go, free to go, meaning they are done. That is we have made a protein, a protein which was in fact the result of transcription followed by translation. As you can imagine, does the mRNA in this case suffer any damage in this process, has translation cut up or altered this in any way. Is it free to do this again?

>> Yes.

>> And again and again, and again, and again, and again, and again. In fact, the interesting question is how long can this

mRNA exist and continue to crank out copies of that same protein? The answer is probably days or weeks, in which case you could produce millions of copies of this particular protein. What ultimately happens is that the mRNA and for that matter the RNA will be degraded, that means fall apart just as result of wear and tear. So is the mRNA you're using tonight going to be used all through the night and probably for the next week or two? Absolutely. But certainly, it's understood that the accuracy of this mRNA determines the accuracy of this protein. And remember, the accuracy of the mRNA which predicated on the intact accuracy of the DNA. So, just to set the stage for an idea you already know about, namely mutation, if something damages a DNA, is that going to lead to faulty RNA? Will that lead to faulty protein? Yeah. And don't get me wrong, protein will still be made but that protein will be useless, in many cases worse than useless. How could a protein made be worse than useless? Remember, what makes a protein useful is the sequence of what? Amino acids. And if one of these amino acids is wrong, well, the protein would be different for sure and therefore, it may not work as an enzyme because its shape is just slightly different. And if doesn't work, why is that worse than useless?

>> It occupies space.

>> Hmm?

>> It occupies space.

>> Well, it occupies space, good. But did we go to a lot of trouble to make this useless molecule? And by trouble, I mean what? We used a lot of energy. We used a lot of amino acids. So that's like assembling a car, and the very end say, "Oops, that one doesn't look right in the dumpster," all right? Worse than what? Worse than useless, because we made a lot of--used a lot effort to get to that point. Now, to just cut this off and finish it for tonight, we said that even the smallest proteins contain, let's say, 50 amino acids. And let's be more realistic and say, "Here, we have a protein that has 900 amino acids." This is just an example, possible? Certainly. And remember, what makes this protein unique is not the fact that it has 900, but the fact that we got amino acids in a certain sequence. A sequence dictated, of course, originally by the DNA molecule. But if you've pieced all these ideas together and if a protein has 900 amino acids, how many basis, minimally, would there have to have been in the DNA gene, which coded for that protein of 900 amino acids?

[Inaudible Remark]

What's that? Twenty-seven hundred. In other words, you took the notion that three times that would be what, 2,700. Good. Because for every amino acid, how many bases are required to code for a given amino acids is three to one, right? And so that's a good thought, but not quite right. And there's a little catch. It's certainly would be at least that, but more or less in fact, way more. Why way more? Well, yeah, you have the start and the stop but that's just six, three at each end, but splicing, right? Great, because, remember, that DNA contains a lot of bogus triplets, which are going to be cut out. And so, because much of the given gene is not exon, but what? Introns. Then, I can't tell you what the number is because this is just hypothetical, but it would be way more than that. And so, if this were an answer you would say way more than 2,700, but certainly 2,700 would not be the answer because of the existence of the introns, which are cut out. Shelby [assumed spelling]?

[Inaudible Remark]

Well, again, every protein molecule is unique, and so there are smallish proteins. An example would be insulin, which is a hormone that, you know, that is less than a hundred amino acids. And then we start getting into mega proteins, and I don't have, you know, an example in mind. But again, it's not so much the size of the molecule that matters, it's the S word, it's the shape or the sequence of the amino acids, and the shape of the molecule. But with that said, if the sequence is different, will the shape of the molecule twist up differently, accordingly? Yes. And so tomorrow's lab indeed, deals with these conceivable events. In other words, we said that DNA was a very stable molecule, remember that remark? Meaning, your DNA today is hopefully the same tomorrow. But are there risk factors? Are there things in our world which can serve to alter our DNA against our will and certainly without our consent or knowledge? Sure. And those factors are called mutagens. And mutagens are in the atmosphere that we survive in, the food that we eat and naturally, in many cases, unavoidable. In many cases, it's very avoidable, but that's another conversation. Still, though, if the DNA

is changed, will the RNA that's formed be the same or different? And if it's different, will the protein be the same or different? And if it's different, is it likely to be better or worst in terms of its performance? More often than not, worse, and that leads to complications in terms of health even basic survival. So, we will be spending a lot of time with those very important ideas of mutation. That's it for the night, though. Quite enough to digest. Again transcription followed by translation.

[Inaudible Remark]

Did you what?

>> Can I get the text answers for the review questions?

>> Yeah. Do you have a flash drive?

>> Yes.

>> All right. We're going to go--well, actually, I don't know if there's a class down there or not, but it's in the lab. So, these are other folks. So, I'm going to walk down there. It's in the lab and we'll copy that for you, if we can.

>> All right.