

DNA:

Structure and Replication

DNA: the stuff of life. Well, not really, despite the hype. DNA does contain the instructions to make a lot of the stuff of life (proteins), although again, not all the stuff of life. At least not directly. Deoxyribonucleic acid (and its very close cousin ribonucleic acid, or RNA) is a very long chain polymer. You may recall that a polymer is just a really big molecule made by connecting many small similar molecules together). In this polymer, the small (monomer) molecules are known as nucleotides, and are composed of a pentose (5-carbon sugar either deoxyribose or ribose), a nitrogenous base, and a phosphate group. As you can see in the figure, the nucleotides only vary slightly, and only in the nitrogenous base. In the case of DNA, those bases are adenine, guanine, cytosine, and thymine. Note the similarity of the shapes of adenine and guanine, and also the similarity between cytosine and thymine. A and G are classified as purines, while C and T are classified as pyrimidines. As long as we're naming things, notice "deoxyribose" and "ribose". As the name implies, deoxyribose is just a ribose without an oxygen. More specifically, where there is a hydroxyl group attached to the 2-carbon of ribose, there is only a hydrogen attached to the 2-carbon of deoxyribose. That is the only difference between the two sugars.

In randomly constructing a single strand of nucleic acid *in vitro*, there are no particular rules regarding the ordering of the nucleotides with respect to their bases. The identities of their nitrogenous bases are irrelevant because the nucleotides are attached by phosphodiester bonds through the phosphate group and the pentose. It is therefore often referred to as the sugar-phosphate backbone. If we break down the word "phosphodiester", we see that it quite handily describes the connection: the sugars are connected by two ester bonds (—O—) with a phosphorous in between. One of the ideas that often confuses students, is the directionality of this bond, and therefore, of nucleic acids in general. For example, when we talk about DNA polymerase, the enzyme that catalyzes the addition of nucleotides in living cells, we say that it works in a 5-prime (5') to 3-prime (3') direction. This may seem like arcane molecular-biologist-speak, but it is actually very simple. Take another look at two of the nucleotides joined together by the phosphodiester bond (fig. 1, bottom left). An adenine nucleotide is joined to a

Using this book: This book is designed to be used in both introductory and advanced cell biology courses. The primary text is generally on the left side of the vertical divider, and printed in black. Details that are usually left to an advanced course are **printed in blue** and found on the right side of the divider. Finally, additional biomedically relevant information can be found **in red print** on either side of the divider.

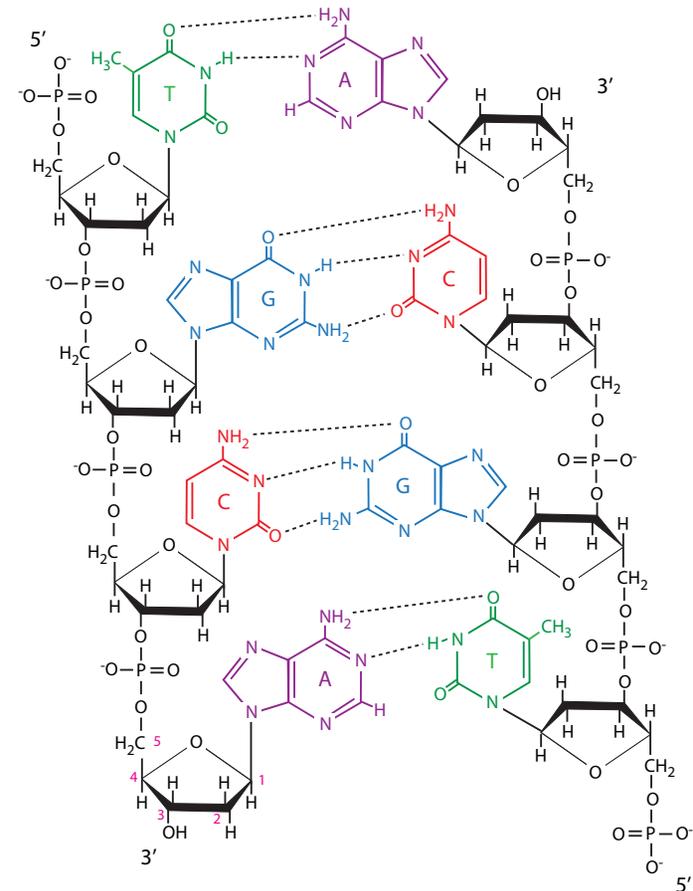


Figure 1. DNA. Deoxyribonucleic acid is a polymer chain of nucleotides connected by 5' to 3' phosphodiester bonds. DNA normally exists as two antiparallel complementary strands held together by hydrogen bonds between adenines (A) and thymines (T), and between guanines (G) and cytosines (C).

cytosine nucleotide. The phosphodiester bond will always link the 5-carbon of one deoxyribose (or ribose in RNA) to the 3-carbon of the next sugar. This also means that on one end of a chain of linked nucleotides, there will be a free 5' phosphate (-PO₄) group, and on the other end, a free 3' hydroxyl (-OH). These define the directionality of a strand of DNA or RNA.

DNA is normally found as a double-stranded molecule in the cell whereas RNA is mostly single-stranded. It is important to understand though, that under the appropriate conditions, DNA could be made single-stranded, and RNA can be double-stranded. In fact, the molecules are so similar that it is even possible to create double-stranded hybrid molecules with one strand of DNA and one of RNA. Interestingly, RNA-RNA double helices and RNA-DNA double helices are actually slightly more stable than the more conventional DNA-DNA double helix.

The basis of the double-stranded nature of DNA, and in fact the basis of nucleic acids as the medium for storage and transfer of genetic information, is *base-pairing*. Base-pairing refers to the formation of hydrogen bonds between adenines and thymines, and between guanines and cytosines. These pairs are significantly more stable than any association formed with the other possible bases. Furthermore, when these base-pair associations form in the context of two strands of nucleic acids, their spacing is also uniform and highly stable. You may recall that hydrogen bonds are relatively weak bonds. However, in the context of DNA, the hydrogen bonding is what makes DNA extremely stable and therefore well suited as a long-term storage medium for genetic information. Since even in simple prokaryotes, DNA double helices are at least thousands of nucleotides long, this means that there are several thousand hydrogen bonds holding the two strands together. Although any individual nucleotide-to-nucleotide hydrogen bonding interaction could easily be temporarily disrupted by a slight increase in temperature, or a miniscule change in the ionic strength of the solution, a full double-helix of DNA requires very high temperatures (generally over 90 °C) to completely denature the double helix into individual strands.

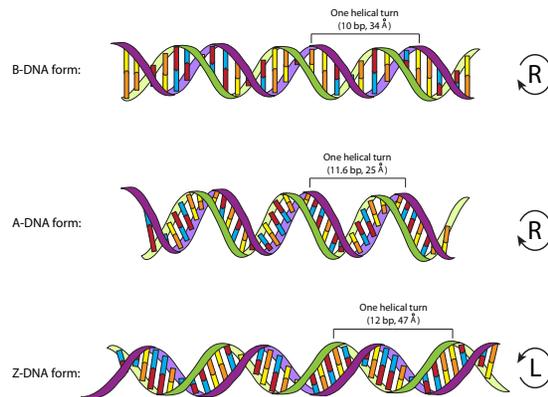
Because there is an exact one-to-one pairing of nucleotides, it turns out that the two strands are essentially backup copies of each other - a safety net in the event that nucleotides are lost from one strand. In fact, even if parts of both strands are damaged, as long as the other strand is intact in the area of damage, then the essential information is still there in the complementary sequence of the opposite strand and can be written into place. Keep in mind though, that while one strand of DNA can thus act as a "backup" of the other, the two strands are not identical - they are complementary. An interesting consequence of this system of complementary and antiparallel strands is that the two strands can each carry unique information.

Bi-directional gene pairs are two genes on opposite strands of DNA, but sharing a promoter, which lies in between them. Since DNA can only be made in one direction, 5' to 3', this bi-directional promoter, often a CpG island (see next chapter), thus sends the RNA polymerase for each gene in opposite physical directions. This has been shown for a number of genes involved in cancers (breast, ovarian), and is a mechanism for coordinating the expression of networks of gene products.

The strands of a DNA double-helix are antiparallel. This means that if we looked at a double-helix of DNA from left to right, one strand would be constructed in the 5' to 3' direction, while the complementary strand is constructed in the 3' to 5' direction. This is important to the function of enzymes that create and repair DNA, as we will be discussing soon. In fig. 1, the left strand is 5' to 3' from top to bottom, and the other is 5' to 3' from bottom to top.

From a physical standpoint, DNA molecules are negatively charged (all those phosphates), and normally a double-helix with a right-handed twist. In this normal (also called the “B” conformation) state, one full twist of the molecule encompasses 11 base pairs, with 0.34 nm between each nucleotide base. Each of the nitrogenous bases are planar, and when paired with the complementary base, forms a flat planar “rung” on

Figure 2. Three conformations of DNA. B-DNA is most common, A-DNA is likely an artifact of crystallization *in vitro*, and Z-DNA may form transiently in parts of the chromosome.



the “ladder” of DNA. These are perpendicular to the longitudinal axis of the DNA. Most of the free-floating DNA in a cell, and most DNA in any aqueous solution of near-physiological osmolarity and pH, is found in this B conformation. However, other conformations have been found, usually under very specific environmental circumstances. A compressed conformation, A-DNA, was observed as an artifact of *in vitro* crystallization, with slightly more bases per turn, shorter turn length, and base-pairs that are not perpendicular to the longitudinal axis. Another, Z-DNA, appears to form transiently in GC-rich stretches of DNA in which, interestingly, the DNA twists the opposite direction.

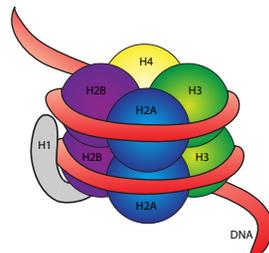
In prokaryotes, the DNA is found in the cytoplasm (rather obvious since there is no other choice in those simple organisms), while in eukaryotes, the DNA is found inside the nucleus. Despite the differences in their locations, the level of protection from external forces, and most of all, their sizes, both prokaryotic and eukaryotic DNA is packaged with proteins that help to organize and stabilize the overall chromosome structure. Relatively little is understood with regard to prokaryotic chromosomal pack-

It has been suggested that both the A and Z forms of DNA are, in fact, physiologically relevant. There is evidence to suggest that the A form may occur in RNA-DNA hybrid double helices as well as when DNA is complexed to some enzymes. The Z conformation may occur in response to methylation of the DNA. Furthermore, the “normal” B-DNA conformation is something of an idealized structure based on being fully hydrated, as is certainly very likely inside a cell. However, that hydration state is constantly changing, albeit minutely, so the DNA conformation will often vary slightly from the B-conformation parameters in figure 2.

aging although there are structural similarities between some of the proteins found in prokaryotic and eukaryotic chromosomes. Therefore, most introductory cell biology courses stick to eukaryotic chromosomal packaging.

Naked DNA, whether prokaryotic or eukaryotic, is an extremely thin strand of material, roughly 11 nm in diameter. However, given the size of eukaryotic genomes, if the DNA was stored that way inside the nucleus, it would become unmanageably tangled. Picture a bucket into which you have tossed a hundred meters of yarn without any attempt whatsoever to organize it by coiling it or bunching it. Now consider whether you would be able to reach into that bucket pull on one strand, and expect to pull up only one strand, or if instead you are likely to pull up at least a small tangle of yarn. The cell does essentially what you would do with the yarn to keep it organized: it is packaged neatly into smaller, manageable skeins. In the case of DNA, each chromosome is looped around a histone complex to form the first order of chromosomal organization: the nucleosome.

Figure 4. The nucleosome is composed of slightly over two turns of DNA around a histone core containing two copies each of H2A, H2B, H3, and H4 histones. The H1 histone is not part of the core unit and functions in coordinating interaction between nucleosomes.



Histones are a family of basic (positively-charged) proteins. They all function primarily in organizing DNA, and the nucleosome is formed when DNA wraps (a little over 2 times) around a core of eight histones - two each of H2A, H2B, H3, and H4. The number and position of the positive charges (mostly from lysines and arginines) are crucial to their ability to tightly bind DNA, which as previously pointed out, is very negatively charged. That “opposites attract” idea is not just a dating tip from the advice columns.

Figure 5. Nucleosomes. (left) Core histone complex - the nitrogens indicative of positively-charged side chains are in blue. (right) DNA wraps around the histone. The red oxygens surrounding the yellow phosphorus atoms are the negatively-charged phosphate groups. Figure from RCSB Protein Data Bank (<http://www.rcsb.org>).

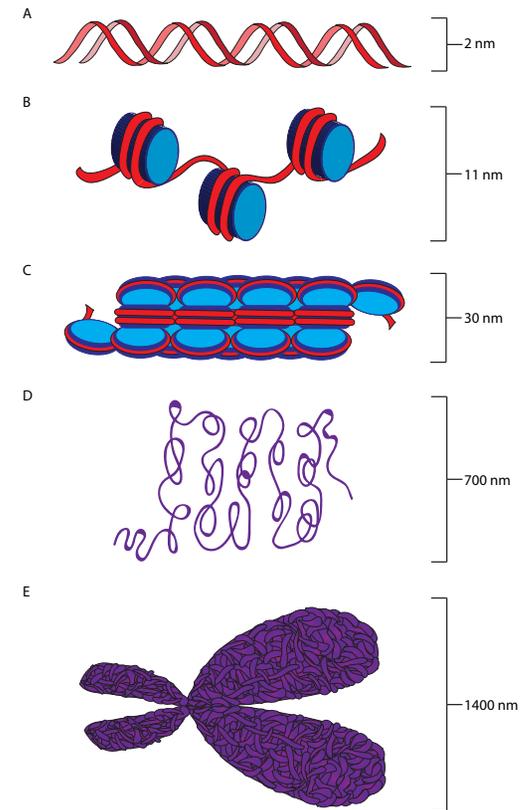
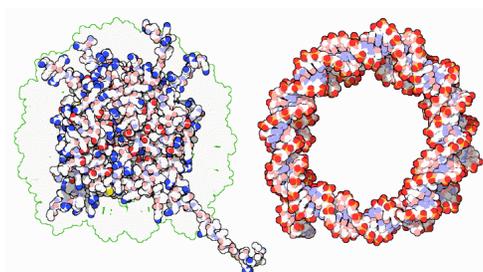


Figure 3. DNA packaging. (A) A naked strand of DNA is approximately 2nm in diameter. (B) Histones, which are octameric proteins depicted here as a roughly cylindrical protein, have positive charges distributed on the outer surface to interact with the negatively-charged DNA backbone. (C) Even the organization afforded by histone binding can leave an unmanageable tangle of DNA, especially with longer eukaryotic genomes, and therefore the histone-bound DNA is packaged into the “30-nm strand”. This is held together, in part, by histone interactions. (D) The 30-nm fibers are looped into 700-nm fibers, which are themselves formed into the typical eukaryotic chromosome (E).

The 30-nm fiber is held together by two sets of interactions. First, the linker histone, H1, brings the nucleosomes together into an approximate 30-nm structure. This structure is then stabilized by disulfide bonds that form between the H2A histone of one nucleosome and the H4 histone of its neighbor.

Upon examination of the 3D structure of the histone core complex, we see that while relatively uncharged protein interaction domains hold the histones together in the center, the positively charged residues are found around the outside of the complex, available to interact with the negatively charged phosphates of DNA.

In a later chapter, we will discuss how enzymes read the DNA to transcribe its information onto smaller, more manageable pieces of RNA. For now, we only need to be aware that at any given time, much of the DNA is packaged tightly away, while some parts of the DNA are not. Because the parts that are available for use can vary depending on what is happening to/in the cell at any given time, the packaging of DNA must be dynamic. There must be a mechanism to quickly loosen the binding of DNA to histones when that DNA is needed for gene expression, and to tighten the binding when it is not. As it turns out, this process involves *acetylation* and *deacetylation* of the histones.

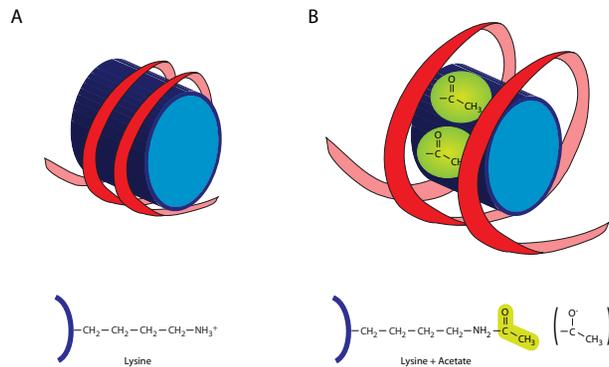


Figure 6. (A) Deacetylated histone allows interaction between the negatively charged phosphates of the DNA and the positively charged lysines of the histone. (B) When the histone is acetylated, not only is the positive charge on the lysine lost, the acetyl group also imparts a negative charge, repelling the DNA phosphates.

Histone Acetyltransferases (HATs) are enzymes that place an acetyl group on a lysine of a histone protein. The acetyl groups are negatively charged, and the acetylation not only adds a negatively charged group, it also removes the positive charge from the lysine. This has the effect of not only neutralizing a point of attraction between the protein and the DNA, but even slightly repelling it (with like charges). On the other side of the mechanism, Histone Deacetylases (HDACs) are enzymes that remove the acetylation, and thereby restore the interaction between histone protein and DNA. Since these are such important enzymes, it stands to reason that they are not allowed to operate willy-nilly on any available histone, and in fact, they are often found in a complex with other proteins that control and coordinate their activation with other processes such as activation of transcription.

Semi-Conservative DNA Replication

DNA replication is similar to transcription in its most general idea: a polymerase enzyme reads a strand of DNA one nucleotide at a time, it takes a random nucleotide from the nucleoplasm, and if it is complementary to the nucleotide in the DNA, the polymerase adds it to the new strand it is creating. Of course, there are significant differences between replication and transcription too, not the least of which is that both strands of DNA are being read simultaneously in order to create two new complementary strands that will eventually result in a complete and nearly perfect copy of an entire organismal genome.

One of the most important concepts of DNA replication is that it is a *semi-conservative* process (fig. 7). This means that every double helix in the new generation of an organism consists of one complete “old” strand and one complete “new” strand wrapped around each other. This is in contrast to the two other possible models of DNA replication, the conservative model, and the dispersive model. A conservative mechanism of replication proposes that the old DNA is used as a template only and is not incorporated into the new double-helix. Thus the new cell has one completely new double-helix and one completely old double-helix. The dispersive model of replication posits a final product in which each double helix of DNA is a mixture of fragments of old and new DNA. In light of current knowledge, it is difficult to imagine a dispersive mechanism, but at the time, there were no mechanistic models at all. The Meselson-Stahl experiments (1958) clearly demonstrated that the mechanism must be semi-conservative, and this was confirmed once the key enzymes were discovered and their mechanisms elucidated.

Prokaryotic Replication

DNA replication begins at an origin of replication. There is only one origin in prokaryotes (in *E. coli*, oriC) and it is characterized by arrays of repeated sequences. These sequences wrap around a DNA-binding protein, and in doing so, exert pressure on the H-bonds between the strands of DNA, and the chromosome begins to unzip in an AT-rich area wrapped around this protein. Remember that A-T pairs are 33% weaker than G-C pairs due to fewer hydrogen bonds. The use of AT-rich stretches of DNA as points of strand separation is a recurring theme through a variety of DNA operations. The separation of the two strands is bidirectional, and DNA polymerases will act in both directions in order to finish the process as quickly as possible. Speed is important here because while replication is happening, the DNA is vulnerable to breakage, and most metabolic processes are shut down to devote the energy to the replication. Even in prokaryotes, where DNA molecules are orders of magnitude smaller than in eukaryotes,

In the Meselson-Stahl experiments, *E. coli* were first incubated with ^{15}N , a heavy isotope of nitrogen. Although it is only a difference in mass of one neutron per atom, there is a great enough difference in mass between heavy nitrogen-containing DNA (in the purine and pyrimidine bases) and light/normal nitrogen-containing DNA that they can be separated from one another by ultracentrifugation through a CsCl concentration gradient (fig. 7).

Over 14 generations, this led to a population of *E. coli* that had heavy nitrogen incorporated into all of the DNA (shown in blue below). Then, the bacteria are grown for one or two divisions in “light” nitrogen, ^{14}N . When the DNA from the bacterial populations was examined by centrifugation, it was found that instead of light DNA and heavy DNA, as would be expected if DNA replication was conservative, there was a single band in an intermediate position on the gradient. This supports a semi-conservative model in which each strand of original DNA not only acts as a template for making new DNA, it is itself incorporated into the new double-helix.

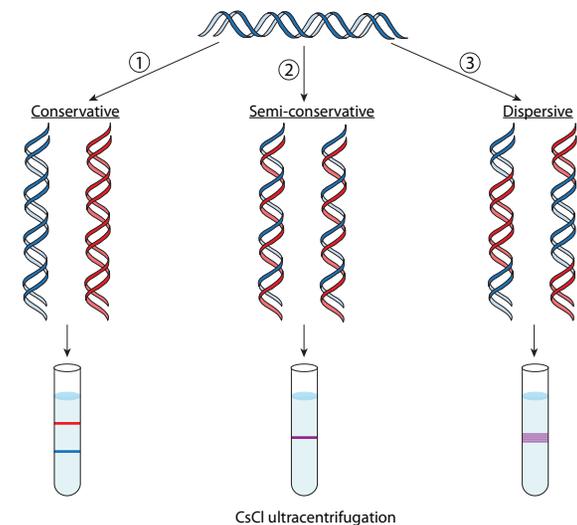


Figure 7. DNA replication. Prior to the discovery of the enzymes involved in replication, three general mechanisms were proposed. In conservative replication, the original DNA strands stay associated with each other, while the newly made DNA forms its own double-helix. Semi-conservative replication posits the creation of hybrid old-new double helices. Dispersive replication proposed molecules composed of randomized fragments of double-old and double-new DNA.

the size of the DNA molecule when it is unraveled from protective packaging proteins makes it highly susceptible to physical damage just from movements of the cell.

The first OriC binding protein, DnaA, binds to DnaA boxes, which are 9 base pair segments with a consensus sequence of TTATCCACA. OriC has five of these repeats, and one DnaA protein binds to each of them. HU and IHF are histone-like proteins that associate with DnaA and together bend that part of the DNA into a circular loop, situating it just over the other major feature of oriC, the 13-bp AT-rich repeats (GATCTNTTNTTTT). DnaA hydrolyzes ATP and breaks the H-bonds between strands in the 13mer repeats, also known as melting the DNA. This allows complexes of DnaB [and DnaC, which is a loading protein that helps attach DnaB(6) to the strand with accompanying hydrolysis of ATP. Also, five more DnaA are recruited to stabilize the loop.] to bind to each single-stranded region of the DNA on opposite sides of the newly opened replication bubble.

DnaB is a helicase; its enzymatic activity is to unzip/unwind the DNA ahead of the DNA polymerase, to give it single-stranded DNA to read and copy. It does so in association with single-stranded-DNA binding proteins (SSBs), and DNA gyrase. The function of SSB is nearly self-explanatory: single-stranded DNA is like RNA in its ability to form complex secondary structures by internal base-pairing, so SSB prevents that. DNA gyrase is a type II topoisomerase, and is tasked with introducing negative supercoiling to the DNA. This is necessary because the unzipping of the DNA by helicase also unwinds it (since it is a double helix) and causes the introduction of positive supercoiling. This means that the entire circular molecule twists on itself: imagine holding a rubber band in two hands and twisting it. As the supercoiling accumulates, the DNA becomes more tightly coiled, to the point that it would be impossible for helicase to unzip it. DnaB/gyrase can relieve this stress by temporarily cutting the double-stranded DNA, passing a loop of the molecule through the gap, and resealing it. This (hopefully) makes a lot more sense

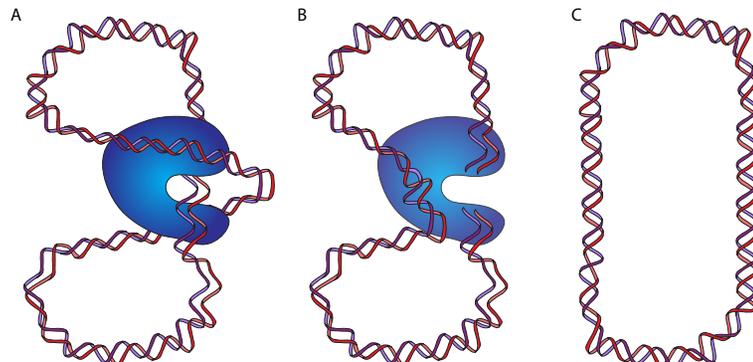
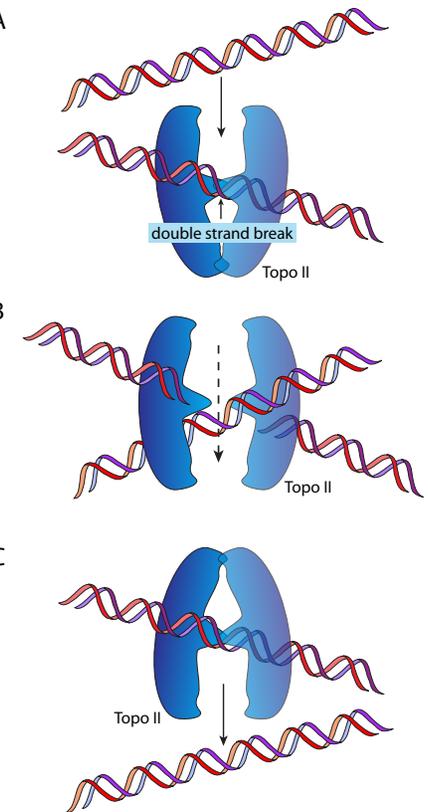


Figure 8. Type II DNA topoisomerases like DNA gyrase relieve supercoiling by making temporary double-strand cut.

Figure 9. Detail of DNA topoisomerase type II action. (A) First, the enzyme binds to the DNA and initiates an endonuclease activity, cutting both strands of the DNA at that point. (B) This complete transection of the DNA (as opposed to the single-strand cut by type I topoisomerases) allows another part of the same DNA molecule to slip through the gap. (C) Finally, the two temporarily broken ends of the DNA, which had been held closely in place by the enzyme.



looking at the diagram. Or, going back to our rubber band, give the rubber band a twist or two, then tape down the two ends. If you snip the rubber band, and pass an adjacent portion of the rubber band through that snip, then reconnect the cut ends, you will find that there is one less twist. Nifty, eh? At this point, some of you are going to say, but if you twist a free-floating rubber band, as one might imagine a free-floating circular DNA chromosome in *E. coli*, you would expect it to naturally untwist. Technically, yes, but due to the large mass of the chromosome, its association with various proteins and the cell membrane, and the viscosity of its environment, it does not behave as though it were completely free.

Once *oriC* has been opened and the helicases have attached to the two sides of the replication fork, the replication machine, aka the replisome can begin to form. However, before the DNA polymerases take positions, they need to be primed. *DNA polymerases are unable to join two individual free nucleotides* together to begin forming a nucleic acid; they can only add onto a pre-existing strand of at least two nucleotides. Therefore, a specialized RNA polymerase (RNAP's do not have this limitation) known as primase is a part of the replisome, and reads creates a short RNA strand termed the *primer* for the DNA polymerase to add onto. Although only a few nucleotides are needed, the prokaryotic primers may be as long as 60 nt depending on the species.

At least five prokaryotic DNA polymerases have been discovered to date. The primary DNA polymerase for replication in *E. coli* is DNA Polymerase III (Pol III). Pol I is also involved in the basic mechanism of DNA replication, primarily to fill in gaps created during lagging strand synthesis (defined 3 pages ahead) or through error-correcting mechanisms. DNA polymerase II and the recently discovered Pol IV and Pol V do not participate in chromosomal replication, but rather are used to synthesize DNA when certain types of repair is needed at other times in the cellular life cycle.

DNA polymerase III is a multi-subunit holoenzyme, with α , ϵ , and θ subunits comprising the core polymerase, and τ , γ , δ , δ' , χ , ψ , and β coming together to form the complete holoenzyme. The core polymerase has two activities: the α subunit is the polymerase function, reading a strand of DNA and synthesizing a complementary strand with great speed, around 150 nt/sec; the ϵ subunit is a 3'-5' "proofreading" exonuclease and acts as an immediate proofreader, removing the last nucleotide if it is incorrect. This proofreading does not reach any further back: it only acts on the most recently added nucleotide to correct misincorporation. Other mechanisms and enzymes are used to correct DNA lesions that arise at other times. [As a matter of nomenclature, *exonucleases* only cut off nucleotides from DNA or RNA from either end, but not in the middle. *Endonucleases* cleave phosphodiester bonds located deeper within a nucleic acid strand.] The θ subunit has no enzymatic activity and regulates the exonuclease function. Although

it has polymerase activity, the Pol III core polymerase has poor processivity - that is, it can only add up to 15 nucleotides before dissociating from the template DNA. Since genomes of *E. coli* strains average near 5 million base pairs, replication in little 15 nt segments would be extraordinarily inefficient.

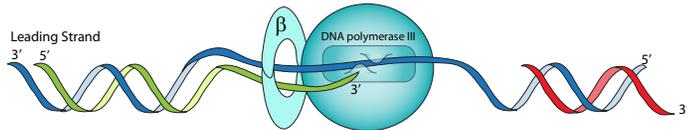


Figure 10. The dimeric β clamp holds DNA Polymerase III on the template, allowing it polymerize more nucleotides before dissociating.

This is where the β subunit is needed. Also known as the β clamp, it is a dimer of semi-circular subunits that has a central hole through which the DNA is threaded. The core polymerase, via an α - β interaction, is attached to this β clamp so that it stays on the DNA longer, increasing the processivity of Pol III to over 5000nt. The β clamp is loaded onto (and unloaded off of) the DNA by a clamp loader complex (also called γ complex) consisting of γ ($\times 3$), δ , δ' , χ , and ψ subunits.

The replication bubble has two replication forks - once the DNA is opened up (unzipped) at the origin, a replication machine can form on each end, with the helicases heading in opposite directions. For simplification, we will consider just one fork – opening left to right – in this discussion with the understanding that the same thing is happening with the other fork, but in the opposite direction.

The first thing to notice when looking at a diagram of a replication fork (fig. 11) is that the two single-stranded portions of template DNA are anti-parallel. This should come as no surprise at this point in the course, but it does introduce an interesting mechanical problem. Helicase opens up the double stranded DNA and leads the rest of the replication machine along. So, in the single-stranded region trailing the helicase, if we look left to right, one template strand is 3' to 5' (in blue), while the other is 5' to 3' (in red). Since we know that nucleic acids are polymerized by adding the 5' phosphate of a new nucleotide to the 3' hydroxyl of the previous nucleotide (5' to 3', in green), this means that one of the strands, called the *leading strand*, is being synthesized in the same direction that the replication machine moves. No problem there.

The other strand is problematic: looked at linearly, the newly synthesized strand would be going 3' to 5' from left to right but DNA polymerases cannot add nucleotides that way. How do cells resolve this problem? A number of possibilities have been proposed,

The clamp loader complex is an ATPase assembly that binds to the β -clamp unit upon binding of ATP (but the ATPase activity is not turned on). When the complex then binds to DNA, it activates the ATPase, and the resulting hydrolysis of ATP leads to conformational changes that open up the clamp temporarily (to encircle or to move off of the DNA strand), and then dissociation of the clamp loader from the clamp assembly.

Understanding the mechanics of DNA replication is a highly visual process, and it is recommended that students frequently flip back and forth between Figure 11 and the text description of replication. In fact, with the extra space around Figure 11, we recommend writing your own description of the process to help understand the mechanism step by step.

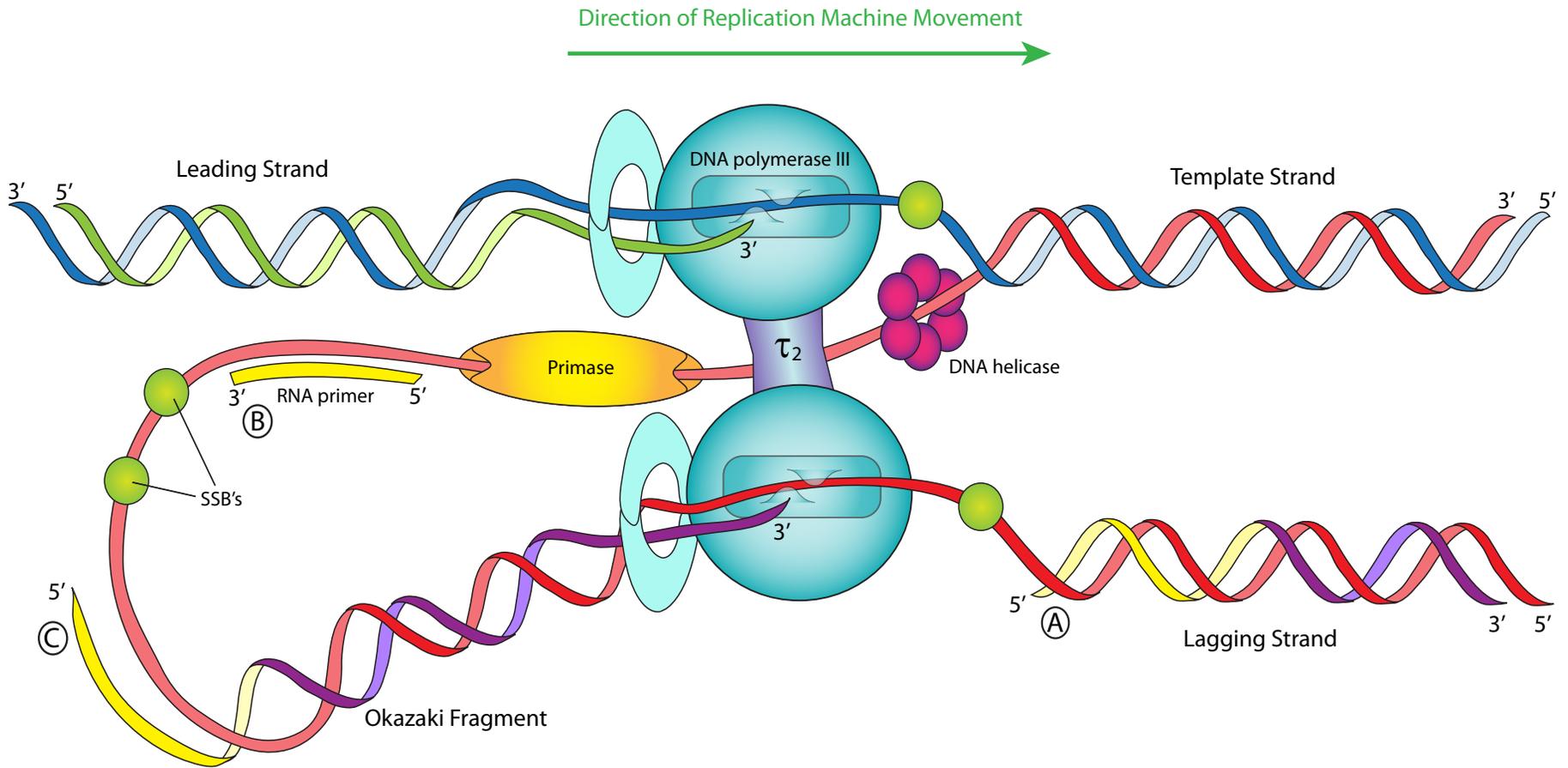


Figure 11. DNA Replication in prokaryotes.

but the current model is depicted here. The replication machine consists of the helicase, primases, and two DNA polymerase III holoenzymes moving in the same physical direction (following the helicase). In fact, the pol III complexes are physically linked through τ subunits.

In order for the template strand that is 5' to 3' from left to right to be replicated, the strand must be fed into the polymerase backwards. This can be accomplished either by turning the polymerase around or by looping the DNA around. As the figure shows, the current model is that the primase is also moving along left to right, so it has just a short time to quickly synthesize a short primer before having to move forward with the replisome and starting up again, leaving intermittent primers in its wake. Because of this, Pol III is forced to synthesize only short fragments of the chromosome at a time, called Okazaki fragments after their discoverer. Pol III begins synthesizing by adding nucleotides onto the 3' end of a primer and continues until it hits the 5' end of the next primer. It does not (and *can not*) connect the strand it is synthesizing with the 5' primer end.

DNA replication is called a semi-discontinuous process because while the leading strand is being synthesized continuously, the lagging strand is synthesized in fragments. This leads to two major problems: first, there are little bits of RNA left behind in the newly made strands (just at the 5' end for the leading strand, in many places for the lagging); and second, Pol III can only add free nucleotides to a fragment of single stranded DNA; it cannot connect another fragment. Therefore, the new "strand" is not whole, but riddled with missing phosphodiester bonds.

The first problem is resolved by DNA polymerase I. Unlike Pol III, Pol I is a monomeric protein and acts alone, without additional proteins. There are also 10-20 times as many Pol I molecules as there are Pol III molecules, since they are needed for so many Okazaki fragments. DNA Polymerase I has three activities: (1) like Pol III, it can synthesize a DNA strand based on a DNA template, (2) also like Pol III, it is a 3'-5' proofreading exonuclease, but unlike Pol III, (3) it is also a 5'-3' exonuclease. The 5'-3' exonuclease activity is crucial in removing the RNA primer (fig. 12). The 5'-3' exonuclease binds to double-stranded DNA that has a single-stranded break in the phosphodiester backbone such as what happens after Okazaki fragments have been synthesized from one primer to the next, but cannot be connected. This 5'-3' exonuclease then removes the RNA primer. The polymerase activity then adds new DNA nucleotides to the upstream Okazaki fragment, filling in the gap created by the removal of the RNA primer. The proofreading exonuclease acts just like it does for Pol III, immediately removing a newly incorporated incorrect nucleotide. After proofreading, the overall error rate of nucleotide incorporation is approximately 1 in 10^7 .

Technically, the 5'-3' exonuclease cleaves the DNA at a double-stranded region downstream of the nick, and may then remove anywhere from 1-10nt at a time. Experimentally, the 5'-3' exonuclease activity can be cleaved from the rest of Pol I by the protease trypsin. This generates the "Klenow fragment" containing the polymerase and 3'-5' proofreading exonuclease.

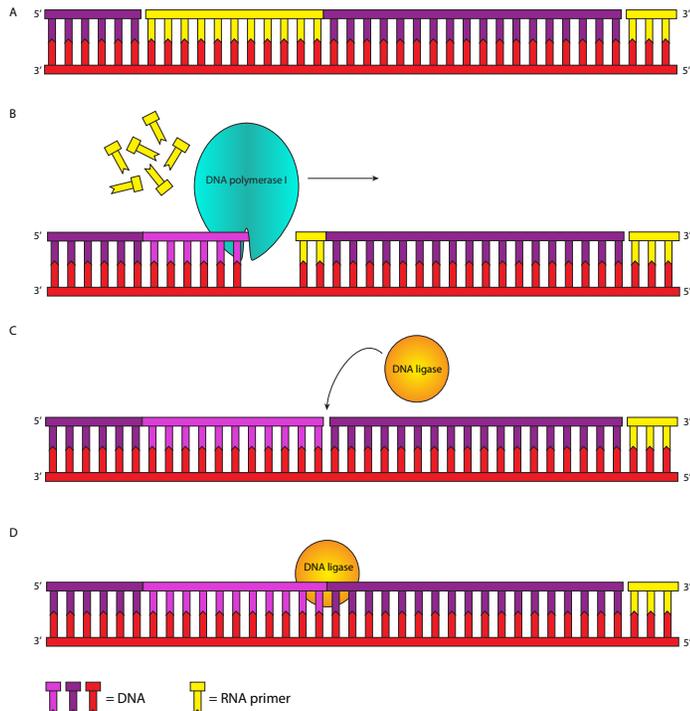


Figure 12. Lagging Strand Synthesis. After DNA polymerase III has extended the primers (yellow), DNA polymerase I removes the primer and replaces it by adding onto the previous fragment. When it finishes removing RNA, and replacing it with DNA, it leaves the DNA with a missing phosphodiester bond between the pol III-synthesized DNA downstream and the pol I-synthesized DNA upstream. This break in the sugar-phosphate backbone is repaired by DNA ligase.

Even though the RNA has been replaced with DNA, this still leaves a fragmented strand. The last major player in the DNA replication story finally appears: DNA ligase. This enzyme has one simple but crucial task: it catalyzes the attack of the 3'-OH from one fragment on the 5' phosphate of the next fragment, generating a phosphodiester bond. This reaction requires energy in the form of hydrolysis of either ATP or NAD⁺ depending on the species (*E. coli* uses NAD⁺) generating AMP and either PP_i or NMN⁺.

Eukaryotic Replication

Given the crucial nature of chromosomal replication for life to exist, it is not surprising to find that eukaryotic DNA replication is very similar to the prokaryotic process. This section will highlight some of the differences, which are generally elaborations on the prokaryotic version.

Unlike prokaryotes, eukaryotic chromosomes often have multiple origins of replication. Considering the size of eukaryotic chromosomes, this is necessary to finish complete replication in a timely manner. Each of these origins defines a *replicon*, or the stretch of the DNA that is replicated from a particular origin. The replicons do not replicate at exactly the same time (although all within the same phase of the cell cycle, see chapter 15), so it is important to make sure that replicons are used only once during a cell cycle.

This requires a “licensing” mechanism. During the cell cycle phase before DNA replication is initiated, a pre-replicative complex is assembled at each origin (Fig. 13). The origins are highly variable in composition and length, ranging from ~100 to well over 10000 base pairs. The pre-RC proteins, on the other hand, are very highly conserved. The pre-RC begins by making the ORC (origin recognition complex, not a creature battling Frodo and Aragorn), which is comprised of six subunits (Orc1-Orc6). Although there is not significant sequence homology, the ORC approximates the function of the DnaA protein in *E. coli*. To complete the pre-RC, the ORC recruits a pair of proteins, Cdc6 and Cdt1 to each side, and they bind the MCM complex (a hexamer of Mcm2-Mcm6 that has an inactive helicase activity), leading to the fully licensed pre-RC. The origin is now ready for activation.

Activation of the pre-RC at an origin of replication requires first Mcm10, which facilitates protein kinases that phosphorylate the MCM complex, activating the helicase activity, and making the replication fork ready to accept the replication machine.

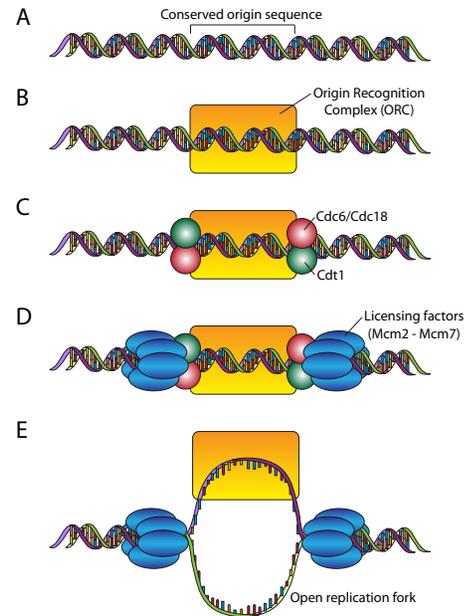


Figure 13. Eukaryotic Origin of Replication

The components of the eukaryotic replication machine may have different names from the prokaryotic, but the functions should be very familiar (Fig. 14). There is a primase to make RNA primers, and tightly associated with the primase is DNA polymerase α . Pol α has neither a 5'-3' exonuclease nor a 3'-5' proofreading exonuclease, but it can synthesize DNA. However, this is not the primary DNA polymerase. The primase/pol α complex is essentially a fancy primase that begins with a short ≈ 10 nt RNA primer and then adds another short ≈ 15 nt DNA fragment, making a *hybrid primer*.

Replication factor C (RFC), acting like the prokaryotic clamp loader complex, then replaces pol α with PCNA, the eukaryotic version of the β clamp. This then recruits DNA polymerase δ , which is the primary replicative DNA polymerase, equivalent to prokaryotic Pol III in function, and necessary for both leading and lagging strand synthesis. Finally, instead of SSB, eukaryotic cells use replication protein A (RPA) to organize and control single-stranded DNA as it is generated during the replicative process.

You may have noticed that none of the eukaryotic DNA polymerases discussed so far have a 5' to 3' exonuclease activity as used by prokaryotic DNA polymerase I to remove primers. In place of that, RNaseH1 and FEN-1 remove the RNA primers (all but one ribonucleotide, and the last ribonucleotide, respectively). Interestingly, FEN-1 also excises chunks of lagging strand DNA within about 15 base pairs of the RNA if they contain mistakes. This seems to help alleviate the problem of lower fidelity of replication by pol α , which has no proofreading capability. After RNaseH1 and FEN-1 have removed primers and near-primer mistakes, pol δ fills in the missing nucleotides, and a ligase enzyme joins the fragments. Pol ϵ does have a 3'-5' exonuclease activity that chops single-stranded DNA into small oligonucleotide fragments and is also associated with the replication machine. The function of Pol ϵ is not clearly understood.

DNA Lesions

The robust nature of DNA due to its complementary double strands has been noted several times already. We now consider in more detail the repair processes that rescue damaged DNA. DNA is not nearly as robust as popular media makes it out to be. In fact, to take the blockbuster book and film, Jurassic Park, as an example, Although there is unquestionably some DNA to be found either embedded in amber-bound parasites, or perhaps in preserved soft tissue (found deep in a fossilized femur, Schweitzer et al, 2007.). It is likely to be heavily degraded, and accurate reproduction is impossible without many samples to work from.

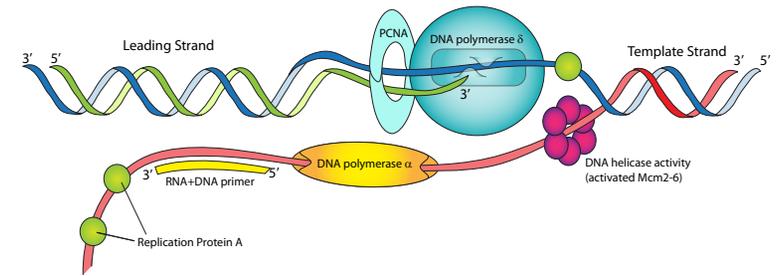


Figure 14. The eukaryotic replication machine is very similar to the prokaryotic machine.

The most common insult to the DNA of living organisms is depurination, in which the β -N-glycosidic bond between an adenine or guanine and the deoxyribose is hydrolyzed. In mammalian cells, it is estimated at nearly 10000 purines per cell generation, and generally, the average rate of loss at physiological pH and ionic strength, and at 37 °C, is approximately 3×10^{-11} /sec. Depyrimidination of cytosine and thymine residues can also occur, but do so at a much slower rate than depurination. Despite the high rate of loss of these bases, they are generally remediated easily by base excision repair (BER), which is discussed later in this section. Therefore it is rare for depurination or depyrimidination to lead to mutation.

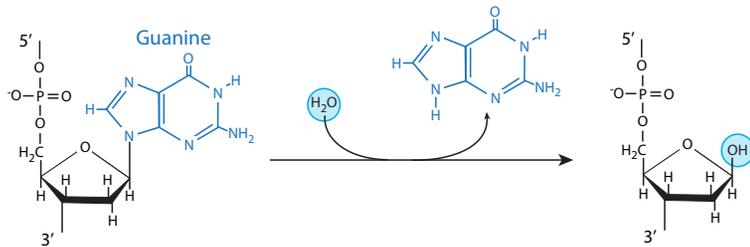


Figure 15. Depurination of guanines (or adenines) is a common DNA lesion.

Three of the four DNA bases, adenine, guanine, and cytosine, contain amine groups that can be lost in a variety of pH and temperature-dependent reactions that convert

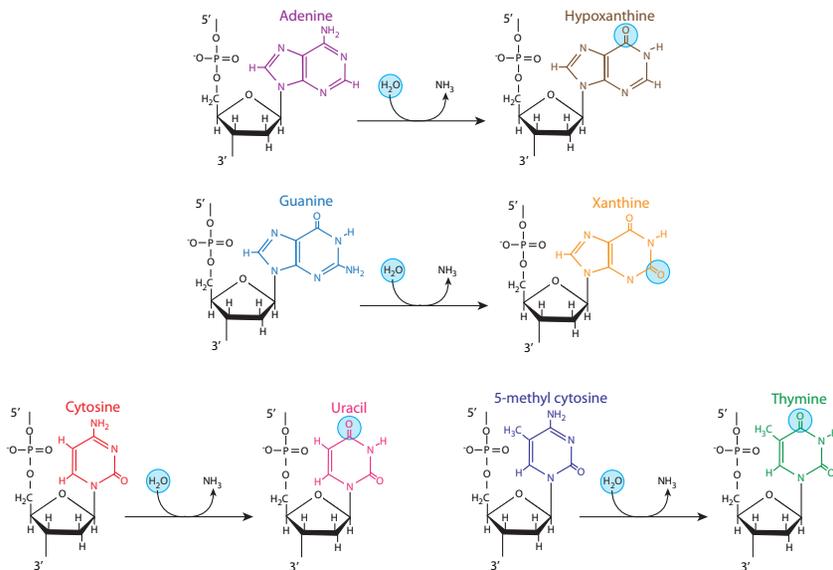


Figure 16. Deamination of adenine, guanine, cytosine, and methylcytosine can lead to mutations upon replication if unrepaired.

Thymine good, Uracil bad. Why is thymine found in DNA rather than uracil? It turns out that the frequency of cytosine deamination may yield a clue as to why cells have gone the extra step (literally, since uracil is a precursor in thymine biosynthesis) to make a new “standard” nucleotide for DNA when uracil worked just fine for RNA, presumably the older genetic molecule. Consider this: if uracil was standard for DNA, then the very frequent deamination conversions of C to U would not be caught by error-checking for non-DNA bases, and the mutation rate would skyrocket. Fortunately, since T has evolved to be the standard base-pairing partner of adenine in DNA, uracil is quickly recognized and removed by multiple uracil DNA glycosylases (more on that later in this chapter), and the integrity of our DNA sequences is much safer.

the bases to hypoxanthine, xanthine, and uracil, respectively. This can sometimes lead to permanent mutations since during replication, they serve as a template for the synthesis of a complementary strand, and where a guanine should go, for example (complementary to cytosine), an adenine may be inserted (because it complements uracil, the deamination product of cytosine). Another deamination, of the modified base methylcytosine, can also lead to a mutation upon replication. Some cytosines may be methylated as part of a regulatory process to inactivate certain genes in eukaryotes, or in prokaryotes as protection against restriction endonucleases. When the methylated cytosine is deaminated, it produces a thymine, which changes the complementary nucleotide (upon replication) from a guanine to an adenine. Deamination of cytosines occurs at nearly the same rate as depurination, but deamination of other bases are not as pervasive: deamination of adenines, for example, is 50 times less likely than deamination of cytosine.

All DNA bases can spontaneously shift to a tautomeric isomer (amino to imino, keto to enol, etc), although equilibrium leans heavily toward one than the other. When a rare tautomer occurs, it base-pairs differently than its more common structural form: guanines with thymines and adenines with cytosines. Here again, a mutation can be propagated during replication of the DNA.

DNA inside a cell must also contend with reactive oxidative species (ROS) generated by the cell's metabolic processes. These include singlet oxygen, peroxide and peroxide radicals, as well as hydroxyl radicals. although it is thought that the hydrogen peroxide and peroxide radicals do not directly attack the DNA but rather generate hydroxyl radicals that do. Most of these ROS are generated in the mitochondria during oxidative phosphorylation and leak out, although some may be generated in peroxisomes, or in some cytosolic reactions. Depending on what part of the DNA is targeted, ROS can cause a range of lesions including strand breaks and removal of bases.

Ionizing radiation (e.g. X-rays) and ultraviolet radiation can each cause DNA lesions. Ionizing radiation is often a cause for double-stranded breaks of the DNA. As described later in the chapter, the repair process for double-stranded breaks necessarily leads to some loss of information, and could potentially knock out a gene. Ultraviolet radiation that hits adjacent thymines can cause them to react and form a cyclobutyl (four carbons bonded in closed loop) thymine dimer. The dimer pulls each thymine towards the other, out of the normal alignment. Depending on the structural form of the dimer, this is sufficient to stymie the replication machine and halt replication. However, some data suggests that normal basepairing to adenine may be possible under some conditions, although, it is likely only one base-pair would result, and the missing base could lead to either random substitution or a deletion in the newly synthesized strand.

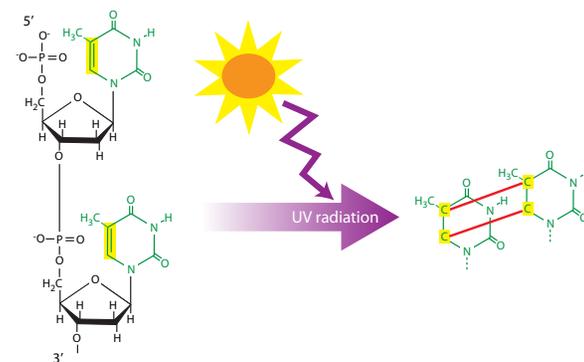


Figure 17. Ultraviolet radiation can be absorbed by some DNA and commonly causes pyrimidine cyclobutyl dimers connecting adjacent nucleotide bases.

Finally, we consider the formation of chemical adducts (covalently attached groups) on DNA. They may come from a variety of sources, including lipid oxidation, cigarette smoke, and fungal toxins. These adducts attach to the DNA in different ways, so there are a variety of different effects from the adducts as well. Some may be very small adducts - many environmental carcinogens are alkylating agents, transferring methyl groups or other small alkyl groups to the DNA. Other adducts are larger, but also attach covalently to a nitrogenous base of DNA. Common examples are benzo(a)pyrene, a major mutagenic component of cigarette smoke, and aflatoxin B1, produced by a variety of *Aspergillus*-family fungi. Benzo(a)pyrene is converted to benzo(a)pyrene diol epoxide, which can then attack the DNA. When this happens, the flat pyrene ring intercalates between bases, causing steric changes that lead to local deformation of the DNA and disruption of normal DNA replication.

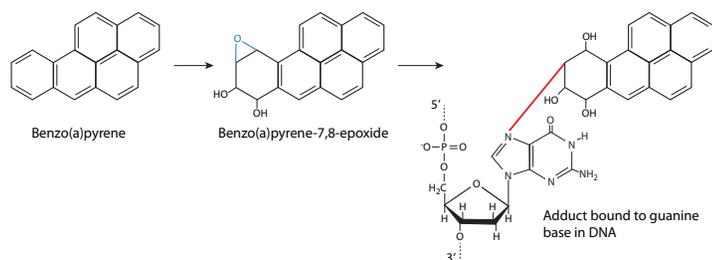


Figure 18. Benzo(a)pyrene is converted to an epoxide form by the cell. The epoxide and form an adduct on DNA.

Aflatoxin B1 is the primary aflatoxin produced by some species (esp. *flavus*, *parasiticus*) of *Aspergillus*, a very common mold that grows on stored grain (as well as detritus and other dead or dying plant matter). In addition to infecting grain, it is a common problem with stored peanuts. At high levels, aflatoxin is acutely toxic, but at lower levels, it has the insidious property of being unnoticeably toxic but mutagenic. Like benzo(a)pyrene, it is metabolized into an epoxide and will then react with DNA to form an adduct that can disrupt replication.

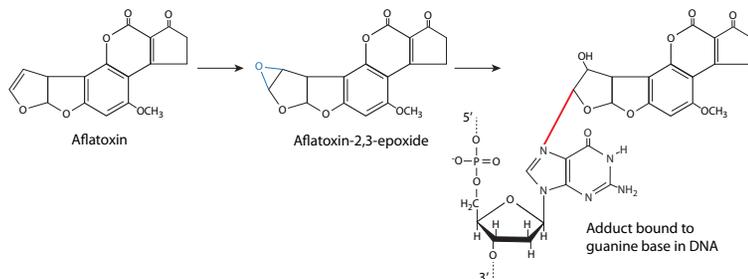


Figure 19. The epoxide form of aflatoxin also forms adducts on DNA.

Some alkylating agents, particularly N-nitroso compounds, are formed in the acidic conditions of the stomach from nitrosation of naturally occurring nitrites produced from food (reduction of nitrates), or environmental nitrites in drinking water. Ironically, while some alkylating agents can cause cancers, others are used therapeutically as anticancer treatments, e.g. mitomycin, melphalan. The idea, as with many cancer treatments, is that although such drugs cause DNA damage to non-cancerous cells as well as cancer cells, the high rate of cancer cell proliferation gives them fewer chances for repair of damaged DNA, and thus greater likelihood that the damage might halt replication and lead to cell death.

In a similar vein, crosslinking chemotherapeutic agents such as cisplatin (a platinum atom bonded to two chloride groups and two amino groups) also bind to DNA. The chloride groups are displaced first by water and then by other groups including sites on DNA. Although sometimes classified as an alkylating agent, it obviously is not, but it acts similarly. Cisplatin goes a step further than a simple alkylating agent though, because it has another reactive site and can thus crosslink (covalently bond) another nucleotide, possibly on another strand of DNA, making a strong obstruction to DNA replication. Cisplatin can also crosslink proteins to DNA.

Benzo(a)pyrene and aflatoxin B1 are not themselves mutagens. Once they are in the cell, the normal metabolism of these compounds leads to diol epoxide formation, which can then attack the DNA. Although the 7-nitrogen (N7) of guanine is more nucleophilic, and is a target for aflatoxin, most benzo(a)pyrene diol epoxide adducts attach to the 2-nitrogen of guanine residues.

There are federal standards (20-300 parts per billion depending on usage) for aflatoxin in various forms of grain-based animal feed, especially corn-based feeds, because the toxin can pass through the animal into milk, as well as linger in the meat. In addition to feed, there are federal maximums for peanuts and peanut products, brazil nuts, pistachios, and other foodstuffs (actionable at 20 ppb).

Well then, what's a poor cell to do when its DNA is being constantly ravaged? As it turns out, there are some very good repair processes that are constantly at work on the DNA, scanning it for defects, and where possible, making repairs. Often the repairs are perfect, if the complementary strand is intact, sometimes mutations must be introduced, and finally there are occasions when repair is impossible, and apoptosis is triggered to kill the cell and prevent propagation of damaged DNA.

DNA Repair

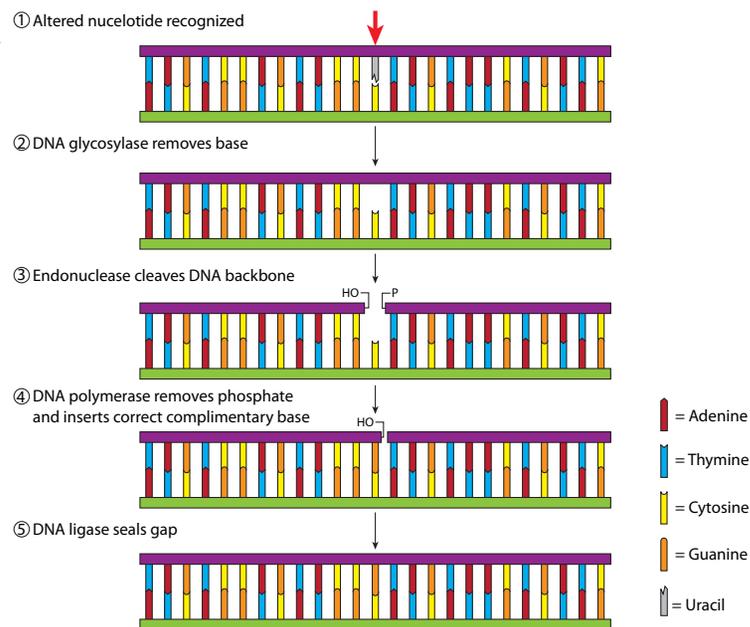
Strictly defined, the simplest repair mechanism does not use an enzyme. Dealkylation, or removal of alkyl groups (like $-\text{CH}_3$ or $-\text{C}_2\text{H}_5$) involves only the transfer of an alkyl group from an O6-methylguanine or O6-ethylguanine onto O6-alkylguanyl-DNA alkyltransferase. Despite the name, the alkyltransferase is not really an enzyme, since it is permanently altered and inactivated by the reaction and therefore does not fit the definition of a catalyst. Note that this does not remediate alkylation at N7 or other sites, just the O6-linked ones.

The next simplest repair mechanism is the uncoupling of pyrimidine cyclobutyl dimers. This can be accomplished through the activity of DNA *photolyases*, also known as photoreactivating enzymes. These are named not just because the formation of the pyrimidine cyclobutyl dimers is usually due to UV light exposure, but because the repair enzymes themselves require exposure to light (300-500nm, near UV to visible blue) to catalyze the dimer-breaking reaction.

While dealkylation and dimer lysis are relatively simple processes that make only a subtle change to the DNA, excision repair mechanisms are more complicated and require multiple enzymatic steps to complete. When a small (not sterically bulky) lesion is limited to a single base, whether missing from depurination or incorrectly formed due to deamination or misincorporation, the process known as *base excision repair* (BER) is engaged. As illustrated in figure 20, if a non-conventional base is recognized, it is then removed by an appropriate *DNA glycosylase*. At present (Genbank search, July 2009), there are at least 8 specific genes encoding human DNA glycosylases, although three encode glycosylases that recognize uracil in various situations. Once the base has been removed by the glycosylase, an endonuclease is enlisted to break the phosphodiester bonds that hold the now-empty phosphodeoxyribose. The resulting gap in the DNA is filled in by a DNA polymerase and finally the strand is reconnected by DNA ligase.

More specifically, DNA photolyases (a ~60kD protein), are non-covalently associated with a chromophore ($\text{N}^5, \text{N}^{10}$ -methylene tetrahydrofolate or 5-deazaflavin) and an FADH^- . The photolyase binds to the pyrimidine cyclobutyl dimer of either single-stranded or double-stranded DNA in a light-independent and sequence-independent manner. However it does not catalyze any change in the bond until light is absorbed by the chromophore, which then transfers the energy to FADH^- , causing it to eject an electron to the dimer, thus breaking it apart.

Figure 20. Base Excision Repair requires five separate steps.



In the case of bulky lesions that significantly alter the physical presentation of the DNA to the polymerases and other enzymes that process DNA, a different type of repair process is involved. Nucleotide excision repair (NER), perhaps better named *polynucleotide* excision repair, involves the removal of the lesion as well as some of the nucleotides in the immediate vicinity. There are two major initiators of NER: either a non-transcriptionally active portion of the DNA is scanned by XPC (fig. 21A), which recognizes a bulky lesion and recruits the repair complex, or as a gene is being transcribed, RNA polymerase runs into a lesion, and then recruits the repair complex via CSA and CSB (fig. 21F and G). If the detection is through XPC, one of the early repair factors recruited to the site is Transcription Factor IIH/XPB/XPD, which is a DNA helicase (fig. 21B). This type of global genome detection is inefficient and relatively slow, but provides a basal level of error-checking for all DNA. In the case of DNA being transcribed, the RNA polymerase complex already includes TFIIH, of which XPB and XPD are a part. This transcriptionally-directed detection is more efficient and targets those parts of the DNA in greatest use in a given cell. In the next step (fig. 21C), XPG, associated with BRCA1/2, and XPF, associated with ERCC1, excise a portion of the affected strand, including but not limited to the lesion itself. DNA polymerase δ or ϵ can then add onto the free 3'OH to fill in the gap based on the complementary strand sequence (fig. 21D). Finally, the repair is connected on its 3' end to the rest of the strand by DNA ligase (fig. 21E).

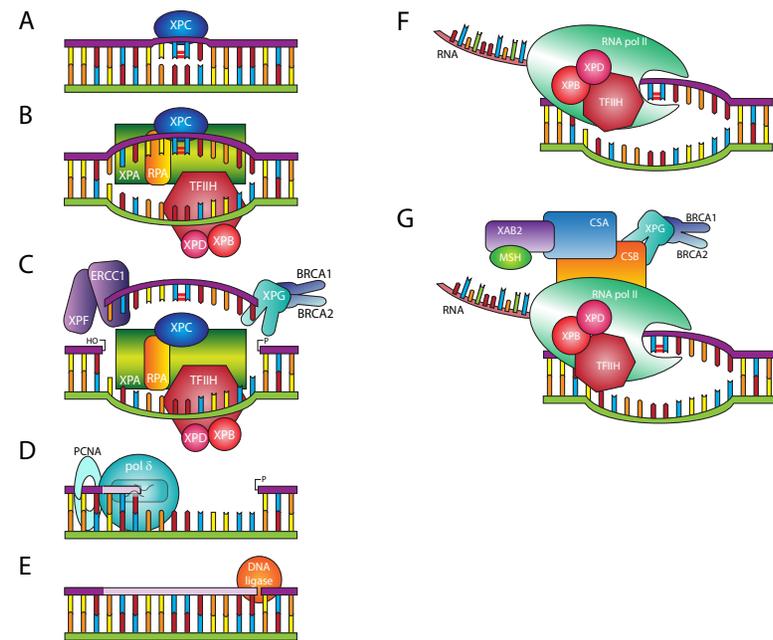


Figure 21. Nucleotide Excision Repair in Eukaryotic Systems

The “XP” in XPC, XPB, XPD, and the others in fig. 18 refers to xeroderma pigmentosa, another autosomal recessive disease, of which the primary characteristic is the formation of skin carcinomas at a young age. Because NER is a major form of pyrimidine dimer repair (in addition to photolyases), its disruption by mutations to one or more of the XP genes leads to extreme sensitivity to UV-induced lesions. Affected individuals must minimize exposure to the sun. The name of the disease comes from the characteristic pigmented lesions (keratoses) that often form on the skin when exposed to sun.

CSA and CSB are named for Cockayne syndrome, an autosomal recessive aging disorder. Mutations in either gene can cause the disorder, which is characterized by premature aging, stunted growth, photosensitivity, and developmental defects of the nervous system. Presumably, knocking out the DNA repair capability of CSA or CSB leads to fast accumulation of damage, inability to transcribe needed genes, and eventually cell death.

A sort of variation on NER is the mismatch repair (MMR) system. This is best understood in prokaryotes: in *E. coli*, MutS is a small protein that forms homodimers at mismatch sites. The MutS dimers recruit two MutL proteins, each of which interacts with one of the MutS units. Each MutS/MutL complex pushes DNA through inwardly, forming a loop with the mismatch in the center of the loop. This continues until one of the MutS/MutL complexes encounters a *hemimethylated* GATC sequence. This causes recruitment of MutH, a highly specialized endonuclease that makes a single-stranded nick in the backbone of the non-methylated strand. This provides an opening for the 3'-5' exonuclease I or the 5'-3' exonuclease VII (or RecJ) to degrade the strand from the nick to the point of mismatch. This is then, as you may have guessed, filled in by DNA polymerase and the backbone connected by ligase. In eukaryotes, multiple homologues to the MutS and MutL proteins have been discovered and the process is similar, but not clearly understood yet, as no homologue to MutH has yet been discovered.

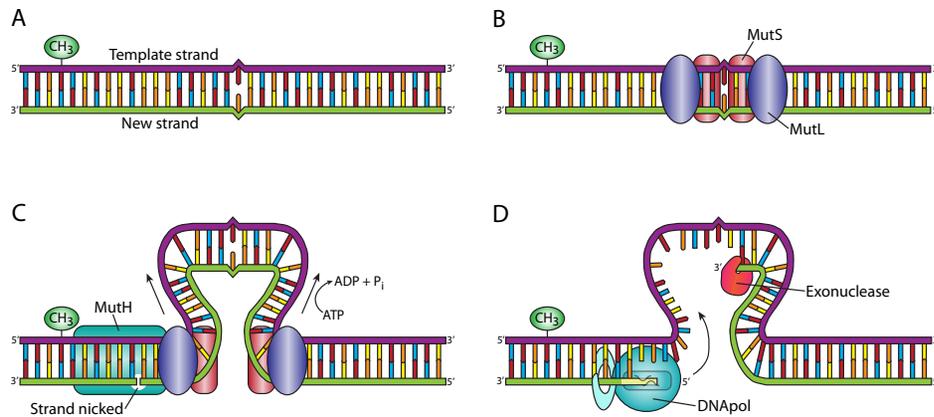


Figure 22. Mismatch Repair System. Mismatched nucleotides do not form normal H-bonds, and cause a minor disturbance (A) in the DNA structure. (B) MutS dimerizes at mismatch sites and each recruits a MutL protein. (C) The MutS/L complex pushes DNA through to form a loop with the mismatch in the center. When a methylated GATC is encountered, the MutS/L stop moving DNA and recruit the MutH endonuclease. An exonuclease then digests the DNA strand to the mismatch. (D) Finally, a DNA polymerase fills in the missing segment and the backbone is sealed by ligase.

So far, the repairs have been based on the assumption that a lesion affects only one strand, and the other strand can provide a reliable template for effecting repairs without the loss of information. Unfortunately, that is not always the case, and some lesions and repair processes necessarily lead to sequence loss. When a double-strand break occurs, perhaps as the product of ionizing radiation, the most common repair mechanism is known as non-homologous end joining (NHEJ). The double-stranded ends are first recognized by Ku, a heterodimeric circular protein that binds the DNA ends. Ku then recruits the kinase DNA-PK_{cs}. The DNA-PK_{cs} acts as a bridge to bring the

Recall that in *E. coli*, Dam methyltransferases eventually methylate the DNA as a method of protecting its genome, but newly synthesized DNA is not methylated. Thus, the assumption is that the methylated strand contains the original and correct base, while the mismatch is due to misincorporation in the newer strand.

Another prokaryotic DNA repair system is the SOS response. As depicted in fig. 23 below, if there is no damage, RecA is inactive, so LexA protein can repress the production of more SOS repair proteins. However, if there is damage, RecA proteins bind to the single stranded DNA and are activated. They in turn cleave the LexA repressor allowing production from a number of DNA repair genes.

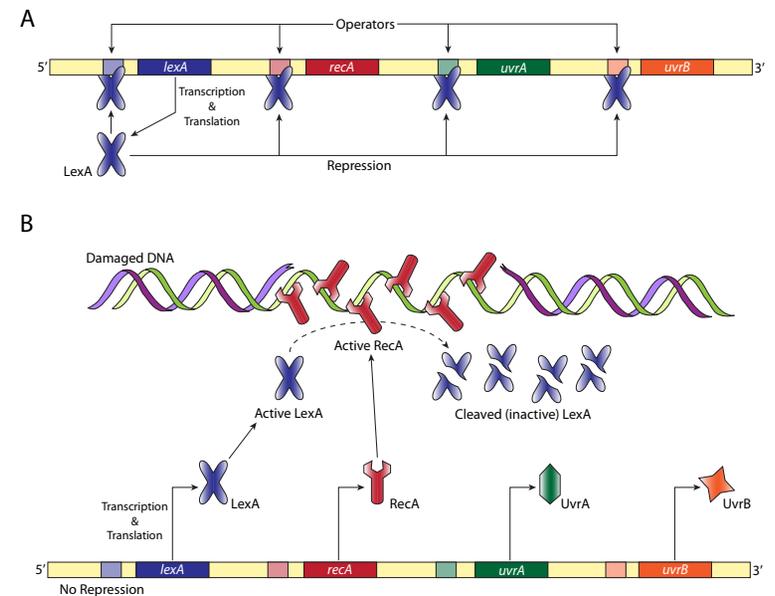


Figure 23. The prokaryotic SOS operon. (A) repressed state when there is no DNA damage, and (B) the activated state in which DNA damage has been detected.

two ends together, and a DNA ligase can then join the ends together. If the strands were broken in different places, resulting in complementary single-stranded overhangs at each end (like those generated by some restriction endonucleases) then the repair is often perfect, since the complementary sequences align the two ends correctly in their original positions. However, if the strand ends have already been acted upon by nucleases and are no longer complementary, then the rejoining of the ends will likely lead to loss of information. In some parts of the DNA, this would have little effect, but if it happened within a gene, the mutated gene product could have abnormal or compromised function.

Telomeres

If there is a mechanism for recognizing loose ends of DNA, what about the ends of every eukaryotic chromosome? They are linear chromosomes, so they have ends, right? What prevents the double-strand-break repair systems from mis-recognizing them all as broken DNA and concatenating all of the chromosomes together? Interestingly, the answer to this question is intimately tied up with the answer to the problem of end-replication, which was very briefly alluded to in our description of replication.

The end-replication problem is one that affects all linear chromosomes. It boils down to one simple fact: an RNA primer is needed to start any DNA replication. So on the 5' end of each strand is an RNA primer (fig. 24 in yellow) that gets removed by the error-correction process. Thus with every round of replication, information is lost from the 5' end of each strand of each chromosome. Eventually, crucial genes are lost and the cell will die; most likely many cellular functions will be compromised long before that happens. The solution to the end-replication problem might be considered more a treatment of symptoms than a cure, to use an analogy to medicine. In short, during the very early stages of an organism's life, a lot of non-coding DNA is added onto the ends of the DNA so that as the cell and its progeny continue to reproduce, the nucleotides do not affect any functional genes. This process is catalyzed by the enzyme, telomerase.

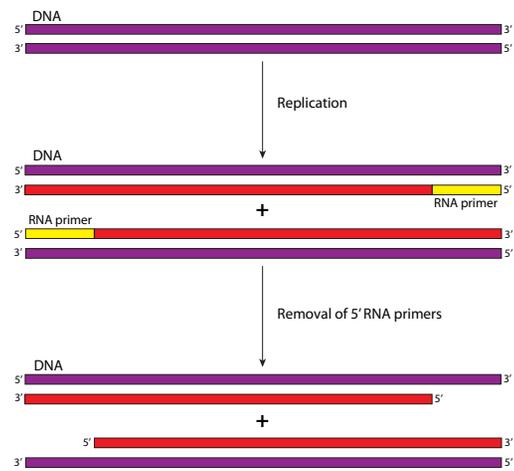


Figure 24. End-replication problem.

breviated TR), and dyskerin. Roughly speaking, the number of telomeric repeats that are placed on a chromosome in early development determines the number of DNA replications and cell divisions that the cell can undergo before succumbing to apoptosis (programmed cell death). Experiments on cells in culture demonstrate a strong correlation between telomere length and longevity, and it is known that cells taken from people with the premature aging disease, progeria, have relatively short telomeres.

Conversely, cancer cells almost universally have upregulated expression of telomerase. Given that a defining characteristic of cancer cells is the ability to proliferate rapidly and indefinitely, turning telomerase back on is, not surprisingly, an important aspect of carcinogenesis. It is therefore a target for anti-cancer treatments; however, to date no telomerase-targeting therapies have proven effective.

Now that we know about telomeres, the question that started this section becomes even more problematic: with these repeated sequence overhangs, how are chromosomes prevented from connecting end-to-end through a double-strand repair-like process? In part due to their repeated sequences, telomeres are able to form end-caps and protect chromosomal ends. The telomeres protect the ends of each chromosome by binding to protective proteins and by forming complex structures. Telomere end binding proteins (TEBP) bind to the 3' overhanging end of the telomere. Other capping proteins, such as the mammalian TRF1 and TRF2 (telomere repeat binding factors) not only bind the telomere, but help to organize it into large looped structures known as T-loops (Fig. 26).

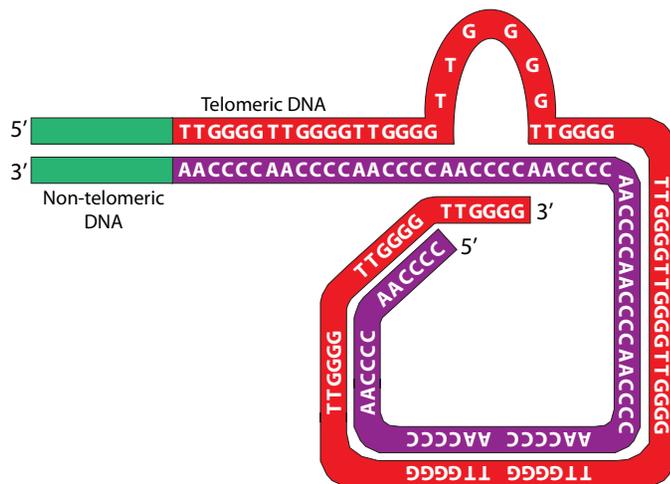


Figure 21. Telomeres form protective T-loop structures on the ends of linear chromosomes.

Finally, the T-loop ends are further stabilized by the formation of G-quartets (fig. 27). G-quartets are a cyclic tetramers that can form in sequences with four consecutive guanine residues, which hydrogen-bond to each other to make a linked square shape stabilized by a metal ion in the center. Furthermore in cases like the telomere, in which such sequences are repeated, the G-quartets can stack and associate three-dimensionally, increasing their stability.

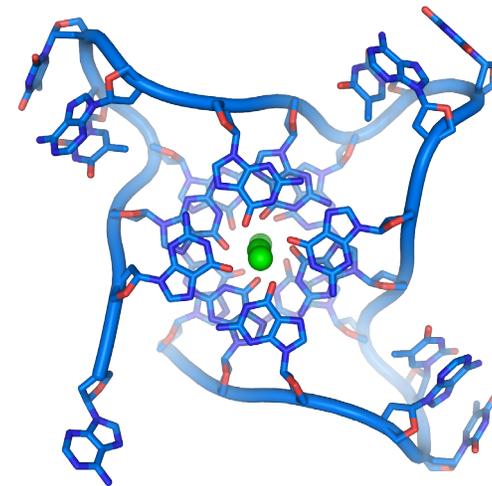


Figure 27. G-quartets form by hydrogen bonding of four consecutive guanine residues stabilized by a metal ion (green), and can form three-dimensional stacks due to telomeric repeats. Produced from Protein DataBank data by T. Splettstoesser.

TRANSCRIPTION :

Reading the Instructions in the Genome

Although DNA is an excellent medium for the storage of information, the very characteristic that makes it so stable and inherently self-correcting - being double-stranded - also makes it unwieldy for using that genetic information to make cell components. Since the informational parts of the molecule (the nitrogenous bases) are locked inside the ladder, reading it requires the energetically expensive task of breaking all the hydrogen bonds holding the two strands together. To do so for every single copy of each protein needed by the cell would not only take a lot of energy, but a lot of time. Instead, there must be a mechanism to take the information from DNA once (or a few times), and then make many copies of a protein from that single piece of information. That mechanism is transcription.

In order to obtain the genetic information in a form that is easily read and then used to synthesize functioning proteins, the DNA must first be transcribed into RNA (ribonucleic acid). As we saw in chapter 1, RNA is extremely similar to DNA, using some of the same nitrogenous bases (adenine, guanine, cytosine) as well as one unique to RNA, uracil. Notice that uracil is very similar to thymine (chapter 7, fig.1), particularly in the placement and spacing of the hydrogen-bonding atoms. Since it is the hydrogen-bonding interaction of these bases (i.e. base-pairing of guanine to cytosine, adenine to thymine/uracil) that forms the basis of information transfer from original DNA to daughter cell DNA, it is logical to expect that the same kind of base-pairing mechanism is used to move the information from a storage state in the double-stranded nucleic acid (DNA) to a more useful/usable state in the form of a single-stranded nucleic acid (RNA).

The process of copying DNA into RNA is called transcription. In both prokaryotes and eukaryotes, transcription requires certain control elements (sequences of nucleotides within the DNA) to proceed properly. These elements are a promoter, a start site, and a stop site. The need for a recognizable point to begin and a point to end the process is fairly obvious. The promoter is somewhat different. The promoter controls the frequency of transcription. If you imagine the needs of a cell at any given time, clearly not all gene products are needed in the same quantity at the same time. There must be a way to control when or if transcription occurs, and at what speed.

Using this book: This book is designed to be used in both introductory and advanced cell biology courses. The primary text is generally on the left side of the vertical divider, and printed in black. Details that are usually left to an advanced course are printed in blue and found on the right side of the divider. Finally, additional biomedically relevant information can be found in red print on either side of the divider.

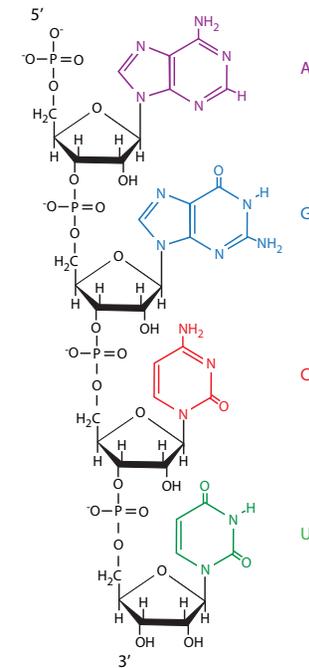


Figure 1. Ribonucleic Acid (RNA) is a polymer of the nucleotides adenine, guanine, cytosine, and uracil, connected by 3'-5' phosphodiester bonds.

The bare-bones version of the process goes something like this: (1) special docking proteins recognize the promoter sequence and bind to it, unzipping a small section around the “start” site; (2) RNA polymerase binds to those special proteins and to the little bit of single-stranded DNA that has just opened up; (3) a helicase enzyme (part of, or attached to the polymerase) unzips the DNA; (4) the RNA polymerase follows behind the helicase, “reading” the DNA sequence, taking ribonucleotides from the environment, matching them against the DNA template, and if they match, adding them to the previous ribonucleotide or RNA chain. This continues until the polymerase reaches the stop site, at which point, it detaches from the template DNA, also releasing the newly made RNA copy of that DNA. Of course, if that was all there was to it, there wouldn’t be entire journals dedicated to studying RNA, its transcription, and the control of that transcription.

The sequence of the promoter is directly related to its function. There may be promoters for housekeeping genes (needed constantly, but at low copy number), “normal” genes (needed as the cell’s situation dictates, rate of transcription also varies), stress response genes (needed rarely), and a variety of other categories. Even within a category, the sequence of the promoter determines its strength. This is based upon what is known as the “consensus sequence”. The consensus sequence is a theoretical “best” promoter based on a survey of all genes in a particular category. The figure below shows an alignment of the promoter sequences of a variety of different genes, all of which are regulated by the same type of promoter and promoter-binding-protein. The highlighted boxes show areas centered around -35 (35 nucleotides upstream of the start site) and -10.

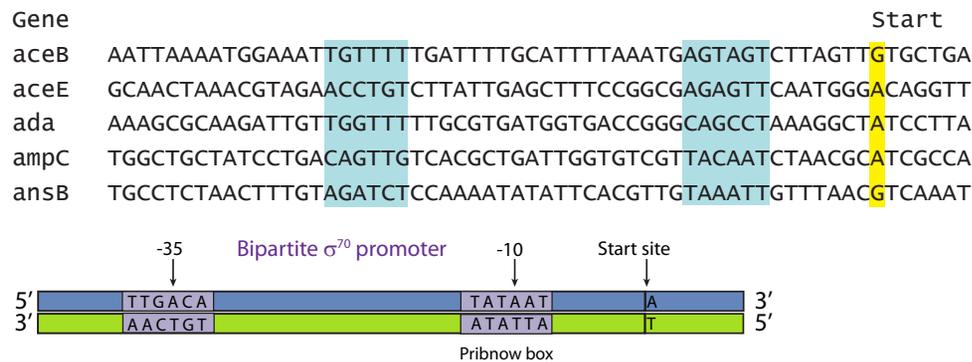


Figure 2. Upstream control sequences of various E. coli genes. Below it is the consensus sequence for σ^{70} promoters.

In contrast to its cellular role as a transient and disposable carrier of genetic information, RNA is thought to have been the primary molecule responsible for making life possible on earth. It has long been postulated that it served dual roles as both a repository of genetic information and as a rudimentary enzyme to act upon that information. Unfortunately, prebiotic chemists have been stymied for decades in coming up with a reasonable synthetic pathway by which RNA could arise from the simple molecules of the earth’s primordial “soup”. The key problem was that ribose could be synthesized, though not particularly efficiently, and bases could be synthesized, but there was no way to connect them together. The chemistry would not allow a condensation reaction between the bases and sugars. In 2009, by leaving behind the conventional idea that ribonucleotides must have been synthesized from ribose and purines/pyrimidines, Powner, Gerland, and Sutherland (*Nature* 459:239-242, 2009) showed that in fact, ribonucleotides could be synthesized from the chemical conditions of a newly formed earth. Rather than attempt to make each “part” and put them together, Powner et al synthesized a molecule that contained parts of what would eventually be both the ribose and a pyrimidine, 2-aminooxazole. Through a series of reactions that utilized phosphate as a catalyst and scavenger, all of which were clearly plausible in the current model of primordial earth, both ribocytidine and ribouridine were created. Of course, this is only the beginning, since this does not extend directly to the formation of purine nucleotides, but it is a very significant step in prebiotic chemistry, and an excellent example of the virtues of “stepping outside of the box” sometimes.

The consensus sequence in fig. 2 shows the most common nucleotide found at each position within those areas of similarity. In this example, the most common prokaryotic promoter is shown: the σ^{70} promoter, so called because it is recognized and bound by the σ^{70} transcription factor. [Here, and by universally accepted convention, recognition sequences of DNA are written as the nucleotides would occur from 5' to 3' on the sense, or non-template, strand.] It is a two-part promoter, with a region centered around -35 (consensus TTGACA), and a region (sometimes called Pribnow box, consensus TATAAT) centered around -10. The (-) sign indicates that the nucleotide is “upstream” of the start site. Upstream means “to the left” when the nucleotides are written as a string of letters, and it means “on the 5' side of” with respect to the 5'-3' directionality of a DNA strand. Notice the relationship between the various individual promoters and the consensus sequence. In general, those promoters with more matches to the consensus sequence are stronger promoters.

A few paragraphs ago, the task of the promoter was defined as controlling the frequency of transcription. How does it do that? What does it mean to be a stronger (or weaker) promoter? First, keep in mind that the expression of any given gene is not automatic, or 100%. At any point in time, many of a cell's genes will be near 0%, or shut off. However, even genes that are turned on are transcribed at different rates. One of the governing factors is the recognition of the promoter site by the RNA Polymerase. For stronger promoters, the RNA polymerase is more likely to recognize the site, dock properly, open up the double helix, and begin transcribing. On the other hand, the RNA polymerase can potentially recognize weaker promoters, but it is less likely to do so, instead passing it by as just another unimportant stretch of DNA. While this is partially a matter of recognition by the polymerase, keep in mind that it is actually governed by recognition of the promoter sequence by the general transcription factors (to be discussed shortly) such as sigma factors in prokaryotes that are recognized by the polymerase.

Notice that there is a high proportion of (A)denines and (T)hymines in the σ^{70} promoter sequences. This is true for many promoters in both prokaryotic and eukaryotic genes. As you probably suspected, this is advantageous because there are only two H-bonds between A-T pairs (as opposed to 3 H-bonds between G-C pairs), which means that it is 33% easier to unzip.

Prokaryotic Transcription

In *E. coli*, as with other prokaryotes, there is only one true RNA polymerase (not including the specialty RNA polymerase, primase, which makes short RNA primers for DNA replication). The polymerase is a multi-subunit holoenzyme comprised primarily of two α subunits, a β subunit, a β' subunit, an ω subunit, and a σ subunit. The α subunits are primarily structural, assembling the holoenzyme and associated regulatory factors. The β subunit contains the polymerase activity that catalyzes the synthesis of RNA, while the β' subunit is used to nonspecifically bind to DNA. The ω subunit is involved in assembly of the holoenzyme and may also play a role in maintaining the structural integrity of the RNA polymerase. Finally, there is the σ subunit, which does not stay closely associated with the core enzyme ($\alpha\beta\beta'\omega$) except when helping to initiate transcription, and is used to recognize the promoter by simultaneously decreasing the affinity of RNAP to DNA in general, but increasing the affinity of RNAP for specific DNA promoter sequences. Why decrease the affinity for non-specific DNA? When the RNAP is not in use, it does not just float about in the nucleoplasm: it is bound quite tightly along the DNA. When the sigma is bound, the decreased affinity allows the RNAP holoenzyme to move along the DNA and scan for promoter sequences. There are multiple isoforms of the σ subunit (such as the sigma-70 mentioned above), each of which recognizes different promoter sequences. All isoforms perform the same basic function of properly locating the RNAP to the start of a gene, and all isoforms only stay attached to the holoenzyme for that one transient purpose, after which they are released (usually after transcribing about ten nucleotides).

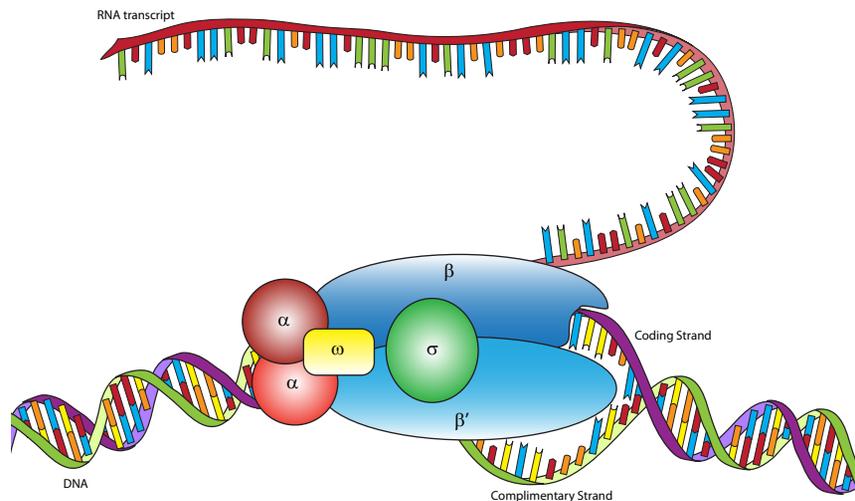


Figure 3. Prokaryotic RNA Polymerases consist of two α subunits, a β , β' , ω , and σ subunits.

Although RNA polymerase was discovered in 1960, the *E. coli* RNAP has not yet been successfully mapped by x-ray crystallography. However, it is very similar to the RNAP of the archaean species, *Thermophilus aquaticus*, which is highly stable (= easier to crystallize) and for which an x-ray crystallographic structure has been elucidated. The data from the Taq RNAP structure and electron microscopic analyses of *E. coli* RNAP produce a lobster-claw-shaped holoenzyme. The inner surface of the claw is lined with positively charged amino acids that can interact with the negatively charged DNA, and when the holoenzyme binds a sigma subunit, the two halves of the claw (formed mostly by the beta and beta' subunits) move closer together to interact with the DNA.

Rifamycins are a class of antibiotics that include rifamycin B, made by the bacteria *Streptomyces mediterranei* (incidentally just one of many antibiotics derived from the *Streptomyces* genus), and rifampicin, its synthetic cousin. They work by binding within the DNA-RNA channel near the active site of RNA polymerase, which sterically prevents the addition of nucleotides to the RNA strand. If the organism cannot transcribe RNA, it cannot use the RNA to make the enzymes and other proteins necessary for life either, and dies. The rifamycin binding site is highly conserved in most prokaryotes but not in eukaryotes, so the antibiotic kills bacteria specifically with little chance of harm to eukaryotes.

Once the holoenzyme has recognized and bound tightly to the DNA at the promoter site, the next step is to “melt” the DNA (breaking the H-bonds and separating the strands of the double helix) in that area so that the RNAP can proceed downstream, read the template DNA strand, and produce the new RNA. Often many RNA transcripts of a gene are needed to produce a large number of active proteins in a short span of time. Highly transcriptionally active genes therefore often have multiple RNA polymerases reading them, one right after another. Generally, an RNA polymerase only needs to process about 15 nucleotides before there is room for another RNAP can bind the promoter and start another transcript.

Strand separation is an energetically difficult process due to the strength of the combined H-bonds between the strands, and often an RNAP may make several short-lived abortive attempts before finally prying open the double helix long and far enough to allow the RNAP to stabilize and transcribe continuously to the stop site.

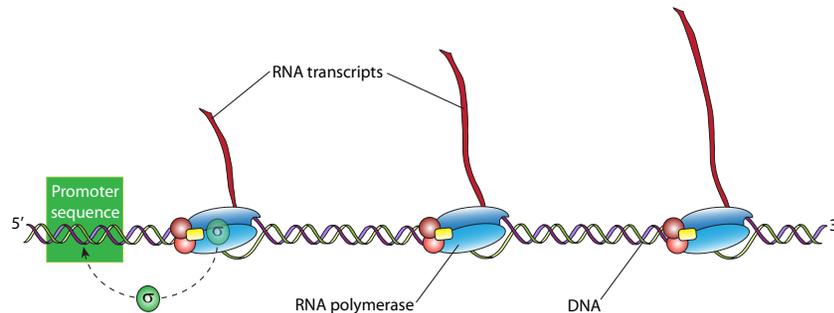


Figure 4. A new RNA polymerase can begin transcribing a gene before the previous one has finished.

The elongation phase of transcription proceeds in a 5' to 3' direction, which is to say that new nucleotides are added to the 3'-OH of the growing strand. Elongation is a stochastic process in which one of the plentiful free-floating ribonucleotides drops into the active site of RNAP opposite the DNA template. If it is the correct nucleotide (complementary to the template), then H-bonds will temporarily form, stabilizing the new nucleotide in place long enough for the RNAP to catalyze the formation of a phosphodiester bond between the 3'-OH of the RNA-in-progress and the 5'-phosphate of the nucleotide. However, if it is the incorrect nucleotide, the proper H-bonds do not form, and the nucleotide usually dissociates from the active site before the RNAP has a chance to bind it to the growing RNA strand. Obviously, this is not a perfect system, and in fact, the error rate for transcription is quite high at approximately 1 in 10000 nucleotides. Fortunately, the cell generally churns out many copies of RNA from any given gene very quickly (approximately 80 nucleotides per second), most of which are either error-free or have errors that do not affect the function of the end-product protein. Furthermore, unlike DNA, in which errors of replication get carried along from one generation of cells to the next, RNA is not a storage medium, and its transient nature means that even mutations that severely impact the protein function only affect the few proteins translated from that one RNA, not the proteins generated from other RNAs made from the same template gene, much less subsequent generations. In other words, to misappropriate a phrase from the movie *Meatballs*, “It just doesn’t matter.”

Eventually, the RNA polymerase reaches the end of the gene and stops transcribing. The termination site is usually marked by a sequence of 4-10 adenine (A) residues on the template strand, and some have a palindromic G-C-rich region that forms a hairpin loop just upstream of the series of adenines. In the first case, it is thought that the resulting string of A-U base pairs is unstable and may lead to the RNAP and the new RNA strand falling off the template DNA while at the same time, the hairpin structure may cause the RNAP to stop or pause, and this can also lead to it dissociating from the DNA. Only about half of all transcription termination sites are marked in this way though, and the others have no significant hairpin loops or easily recognizable sequences other than a series of G-C-rich regions. In this type of termination site, the enzymatic co-factor, rho, is required for termination, and so this is known as rho-dependent termination. Rho is an RNA-binding protein with helicase activity, so it is postulated that it effects termination by forcing the RNA strand off of the DNA template.

Eukaryotic Transcription

Transcription in eukaryotes is more complicated, but follows the same general ideas. The promoter sequences are much more varied both in placement (with respect to the start site) and size. As we will see in the next chapter, eukaryotic genes have many more control elements regulating their expression than do prokaryotic genes. Not only are there more control elements, there are also more RNA polymerases, which serve different specific cellular functions. Obviously, the broad function and location of all the RNA polymerases is the same: read a DNA template and transcribe an RNA copy of it; and since the DNA is found only in the nucleus, so are the polymerases. However, the polymerases differ in exactly what kinds of RNA they produce. RNA Polymerase I is specialized for producing pre-rRNA (rRNA = ribosomal RNA). The pre-rRNA is cleaved post-transcriptionally and incorporated into the ribosomes. Since ribosomes are assembled in the the nucleolus, that is the part of the nucleus in which most RNA Polymerase I is concentrated. RNA Polymerase III also makes an RNA (5S) that is incorporated into the ribosome. It also makes other *untranslated* RNAs such as tRNAs and a variety of small nuclear RNAs. The only RNA polymerase that makes the translatable RNA (mRNA, or messenger RNA) that most people think of when RNA is referred to generically, is RNA polymerase II. This is the RNA polymerase that produces pre-mRNA, which after some processing, becomes mRNA, is transported out of the nucleus, and finally translated into proteins. All of the eukaryotic RNA polymerases are composed of two large subunits, roughly analogous to the β and β' subunits of prokaryotic RNAP, but instead of just three or four other subunits, there are over a dozen smaller subunits to the eukaryotic RNA polymerase holoenzymes.

Initiation of transcription is also much more complicated. Not only is there great variety in promoters recognized by RNAP II, both RNAP I and RNAP III recognize promoters with particular structural characteristics. One of the most common eukaryotic RNAP II promoters is the TATA box, named for the highly conserved motif that defines it. Although it appears similar to the Pribnow box in prokaryotes, it is generally located further upstream from the start site, and its position is far more variable. Whereas the Pribnow box is located at -10, the TATA box may be located closer to -30 +/- 4. Also, rather than just a sigma factor to recognize the promoter in conjunction with the polymerase core enzyme, the eukaryotic promoter is recognized by a multi-subunit complex called transcription factor IID (TFIID). TFIID is comprised of TATA-binding protein (TBP)

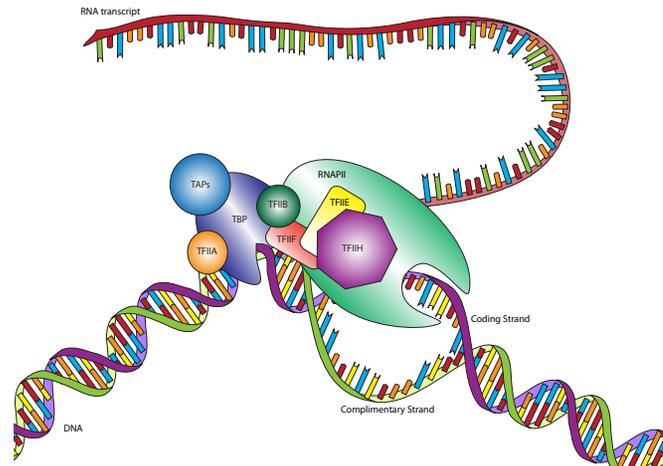


Figure 5. Eukaryotic Transcription. An initiation complex of several transcription factors is needed to dock the RNA Polymerase II in position to begin transcription.

and several TBP-associated factors (TAFs). This binding of the promoter by TFIID occurs independently of RNA Polymerase II, and in fact, RNAP II will not attach to TFIID at this time. After TFIID has bound the TATA box, two more transcription factors, TFIIA and TFIIB, attach to the TFIID as well as the nearby DNA, stabilizing the complex. TFIIF attaches to TFIID and TFIIB to allow docking of the RNA Polymerase II. The complex is still not ready to begin transcription: two more factors are required. TFIIE binds TFIIF and RNAP II, and finally, TFIIH attaches to RNAP II, providing a helicase activity needed to pry apart the two strands of DNA and allow the polymerase to read one of them. TFIIH also has another important enzymatic activity: it is also a serine kinase that phosphorylates the carboxyl-terminal domain (CTD) of RNA polymerase II. There are several serines in the CTD, and as they are sequentially phosphorylated, the CTD extends like a (negatively charged) tail and helps to promote separation between the RNAP II and the TFIID/promoter.

The eukaryotic RNA polymerases were named I, II, and III based on their elution order from ion-exchange chromatography purification. They are also partially distinguishable by their sensitivity to α -amanitin and related amatoxin-family mushroom poisons. RNAP I (and prokaryotic RNAP) is insensitive to these toxins, RNAP III is somewhat sensitive (K_d -10^{-6} M), and RNAP II is highly sensitive (K_d -10^{-8} M). These toxins act by binding to a site in the RNA-DNA cleft and interfering with translocation of the RNA. That is, there is no problem with importing a nucleotide or with attaching it to the new RNA, but the RNA strand cannot move through the active site and allow the next nucleotide to be added.

Elongation of the RNA strand in eukaryotes is very similar to that in prokaryotes with the obvious difference that transcription occurs in the nucleus rather than in the cytoplasm. Thus, in prokaryotes, the RNA can be used for translation of proteins even as it is still being transcribed from the DNA! In eukaryotes, the situation is significantly more complex: there are a number of post-transcriptional events (5' end-capping, 3' polyadenylation, and often RNA splicing) that must occur before the RNA is ready to be transported out of the nucleus and made available for translation in the cytoplasm.

Termination of eukaryotic transcription is not well-described at this writing. RNAP I appears to require a DNA-binding termination factor, which is not analogous to the prokaryotic Rho factor, which is an RNA binding protein. RNAP III terminates transcription without any external factor, and this termination usually occurs after adding a series of uridine residues. However, it does not appear to use the hairpin loop structure found in rho-independent bacterial transcription. The termination of protein-coding RNAP II transcripts is linked to an enzyme complex that also cleaves part of the 3' end of the RNA off, and adds a poly-A tail. However, it is not clear how the polyadenylation complex is involved in determining the point of transcription termination, which can be over 1000 nucleotides beyond the poly-A site (e.g. the β -globin gene in *Mus musculus*). Upon termination and release from the RNAP II and template DNA, the RNA is known as the primary transcript, but must undergo post-transcriptional processing before it is a mature messenger RNA (mRNA) ready to be exported to the cytoplasm and used to direct translation.

Post-Transcriptional Processing of RNA

The first of the post-transcriptional events is 5' end capping. Once the 5' end of a nascent RNA extends free of the RNAP II approximately 20-30 nt, it is ready to be capped by a 7-methylguanosine structure. This 5' "cap" serves as a recognition site for transport of the completed mRNA out of the nucleus and into the cytoplasm.

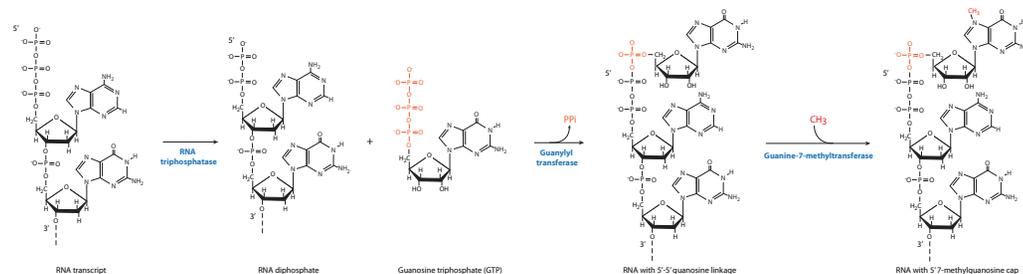


Figure 6. Capping the 5' end of eukaryotic transcripts.

The process actually involves three steps. First, RNA triphosphatase removes the 5'-terminal triphosphate group. Guanylation by GTP is catalyzed by capping enzyme, forming an unusual 5'-5' "backward" bond between the new guanine and the first nucleotide of the RNA transcript. Finally, guanine-7-methyltransferase methylates the newly attached guanine.

On the opposite end of the RNA, on the free 3'-OH, *polyadenylation* occurs. As noted previously, an enzyme complex that docks to a site on the CTD tail of RNAP II cleaves a portion of the 3' end near an AAUAAA recognition sequence and then serially adds a large number of adenine residues. The poly(A) tail is not required for translation, but it has an effect on the stability of transcripts in the cytoplasm. As mRNA molecules stay in the cytoplasm longer, the poly(A) tail is gradually removed. Once the poly(A) tail is gone, the mRNA will soon be destroyed. mRNA molecules with longer poly(A) tails are generally longer-lived in the cytoplasm than those with shorter tails, but there is currently no evidence for a directly proportional effect.

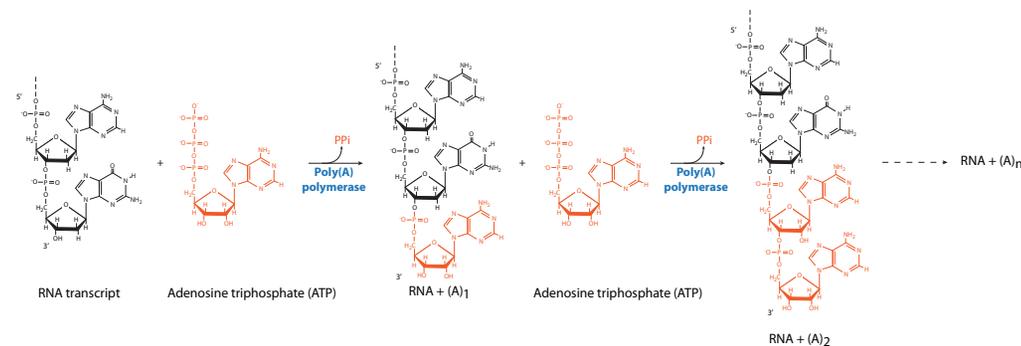


Figure 7. Polyadenylation of the primary transcript.

The third and most complicated modification to newly-transcribed eukaryotic RNA is *splicing*. Unlike prokaryotic RNA, which is a continuously translatable coding region immediately as it comes out of the RNA polymerase, most eukaryotic RNAs have interrupted coding regions. Splicing is the process by which the non-coding regions, known as *introns*, are removed, and the coding regions, known as *exons*, are connected together. In some RNAs, this can happen autonomously, with part of the RNA acting as an enzymatic catalyst for the process. This requires that the RNA have a specific secondary and tertiary structure, bringing the two exons close together while looping out the intron. It was the study of this phenomenon that led to the discovery of ribozymes, which are enzymes made of RNA.

In most cases, however, splicing is carried out by a multi-subunit protein complex known as the spliceosome. Whether it is self-spliced or by spliceosome, there are three main sequence components needed to define an intron that is going to be spliced out (fig. 8). There is a 5' splice site with the consensus sequence AG|GUAAGU. There is a 3' splice site that starts with an 11-nucleotide polypyrimidine tract followed by NCAG|G. And somewhere in between the two, there is a branchpoint adenine, typically within a YNCURAY sequence (Y is a pyrimidine, N is any nucleotide, R is a purine). Splicing is actually a set of two sequential transesterification reactions, and requires physical prox-

Although the enzyme that cleaves the primary transcript in preparation for polyadenylation has not been identified, two non-enzymatic factors, the excitingly-named cleavage factor I (CFI) and cleavage factor II (CFII) have been implicated. The serial adenylation comes from the activity of poly(A) polymerase (PAP) in conjunction with CPSF (cleavage and polyadenylation specificity factor), which binds to the RNA. PAP itself has relatively poor affinity for RNA. As with other nucleic acid polymerases, it adds new nucleotides onto the free 3'-OH of the pre-existing chain. To encourage processivity (continuous polymerization) poly(A) binding protein II (PABII) joins the polyadenylation complex, and is involved in controlling the final length of the poly(A) tail. It should be noted that PABII is a nuclear protein and should not be confused with PABP (poly(A) binding protein) which binds to mRNA molecules in the cytoplasm and plays a role in protecting them from nuclease attack.

Until the discovery of ribozymes, it had been assumed that only the enzymes could only be generated with the diversity of structures possible with the amino acids in proteins.

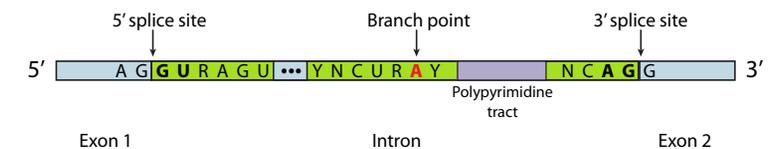


Figure 8. Consensus sequences for splicing.

imity of the reactive sites by bending and looping of the RNA, either autonomously or around protein factors known as snRNPs (pronounced “snurps”). SnRNPs is an acronym for small nuclear ribonucleoproteins. They contain both a protein and a small nuclear RNA (snRNA) component; the latter helps with sequence recognition. Examination of the structure of the snRNA part of these spliceosome snRNPs shows that they are very similar to the shapes taken by the RNA transcript itself in cases of self-splicing. Keeping that in mind, much of the following description of spliceosome-mediated splicing happens in self-splicing as well.

Although the snRNPs are the primary components of the spliceosome, a variety of other splicing factors also play a role. The most prominent are U2AF (U2-associated factor, which binds to the polypyrimidine tract, and SF1 (splicing factor 1, aka branch-point protein BPP) which binds to consensus sequence near the branchpoint. Together they help to properly position the U2 snRNP. There are also a variety of other less-studied splicing factors from the SR protein family (C-terminal Serine-Arginine binding motif) and the hnRNP (heterogenous nuclear ribonucleo-protein) families that act to recruit the primary members of the spliceosome to their proper locations.

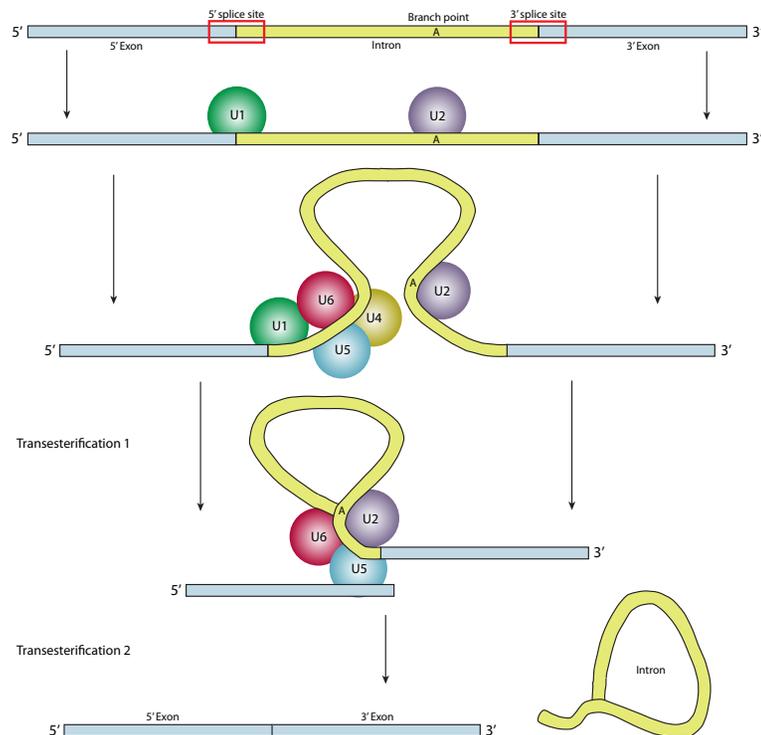


Figure 9. RNA splicing by spliceosome. Description in text below.

In the first step, the U1 snRNP binds to the intronic portion of the 5' splice site. Next, the U2 snRNP binds to the consensus site around the branchpoint, but importantly, there is no base-pairing to the branchpoint A itself. Instead, due to basepairing of U2 with the surrounding sequence, the branchpoint A is forced to bulge out from the rest of the RNA in that region. U4, U5, and U6 join the spliceosome together, but while U5 binds to the 5' exon, and U6 displaces U1 at the 5' splice site, U4 is only transiently attached and also falls off the spliceosome before the first transesterification reaction. As the figure shows, in this reaction, the 2'-OH of the branchpoint A nucleophilically at-

tacks the 5'-phosphate of the first intron nucleotide to form a lariat structure in which the 5' end of the intron is connected to the branchpoint via a 2',5'-phosphodiester bond. This releases the 5' exon (and the whole 5' half of the RNA for that matter), but it is kept in close proximity to the 3' exon (and the rest of the RNA) by U5, which attaches to both exons. This allows the second transesterification to take place, in which the 3'-OH of the first exon attacks the 5' phosphate at the beginning of the second exon, thus simultaneously breaking the bond between the intron and the second exon, and also connecting the two exons via a conventional 3',5'-phosphodiester bond. The intron, in the shape of a lariat, is thus released and will be quickly degraded.

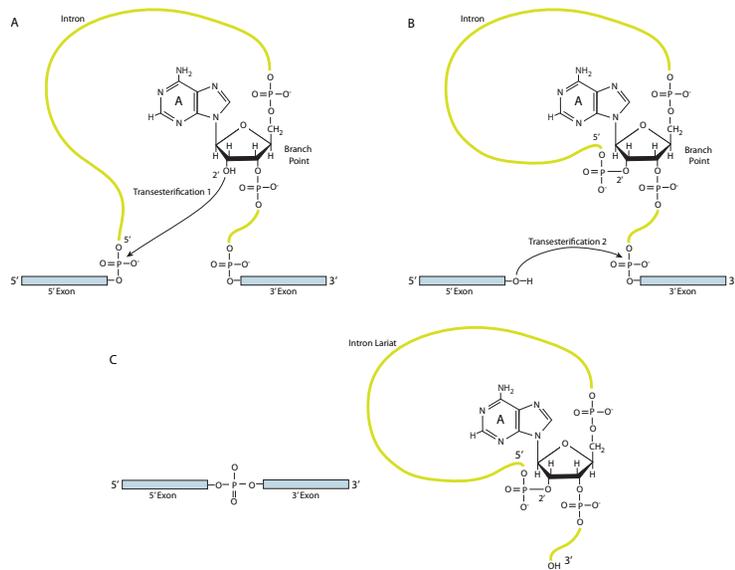


Figure 10. (A) Transesterification reaction 1 connects the 2'-OH of the branchpoint ribose to the 5'-phosphate of the 3' end of exon 1. (B) Transesterification 2 is an attack by the remaining -OH on the 3' end of exon 1 on the 5' phosphate of exon 2. (C) This simultaneously releases the lariat and connects the exons.

Splicing is an efficient (with respect to genome size) way to generate protein diversity. In alternative splicing, some potential introns may be spliced out under certain circumstances but remain as coding sequence under other circumstances. Recall that the splice sites are recognized by base-pairing and therefore, there can be stronger and weaker splice sites depending on how close they are to the consensus and the complementary sequence on the snRNPs. Therefore, a gene with several potential introns may have all introns spliced out 80% of the time, but the other 20% of the time, perhaps only one or two introns are spliced out. Adding variability, there are splicing factors that may bind near splice sites and can either make them more easily recognizable, or nearly hidden.

The classic example of alternative splicing is the gene encoding α -tropomyosin (fig. 11). By splicing in/out different combinations of exons, a single gene can generate seven different proteins, depending on the tissue type. In these cases, particular types of cells or tissues contain specific combinations of splicing factors, and therefore control the recognition of specific splice sites, leading to the different splicing patterns.

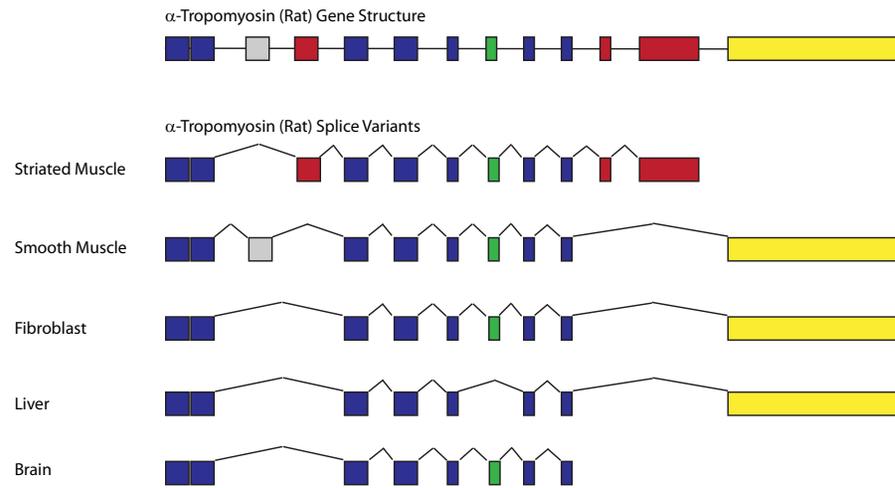


Figure 11. Alternative splicing of the α -tropomyosin gene leads to different forms of the mRNA and protein in different cell types.

Although this concludes the discussion of basic mechanisms of transcription, the next chapter is really a continuation of this one: control of gene expression in its simplest form is regulating the recognition of a promoter sequence by an RNA polymerase.

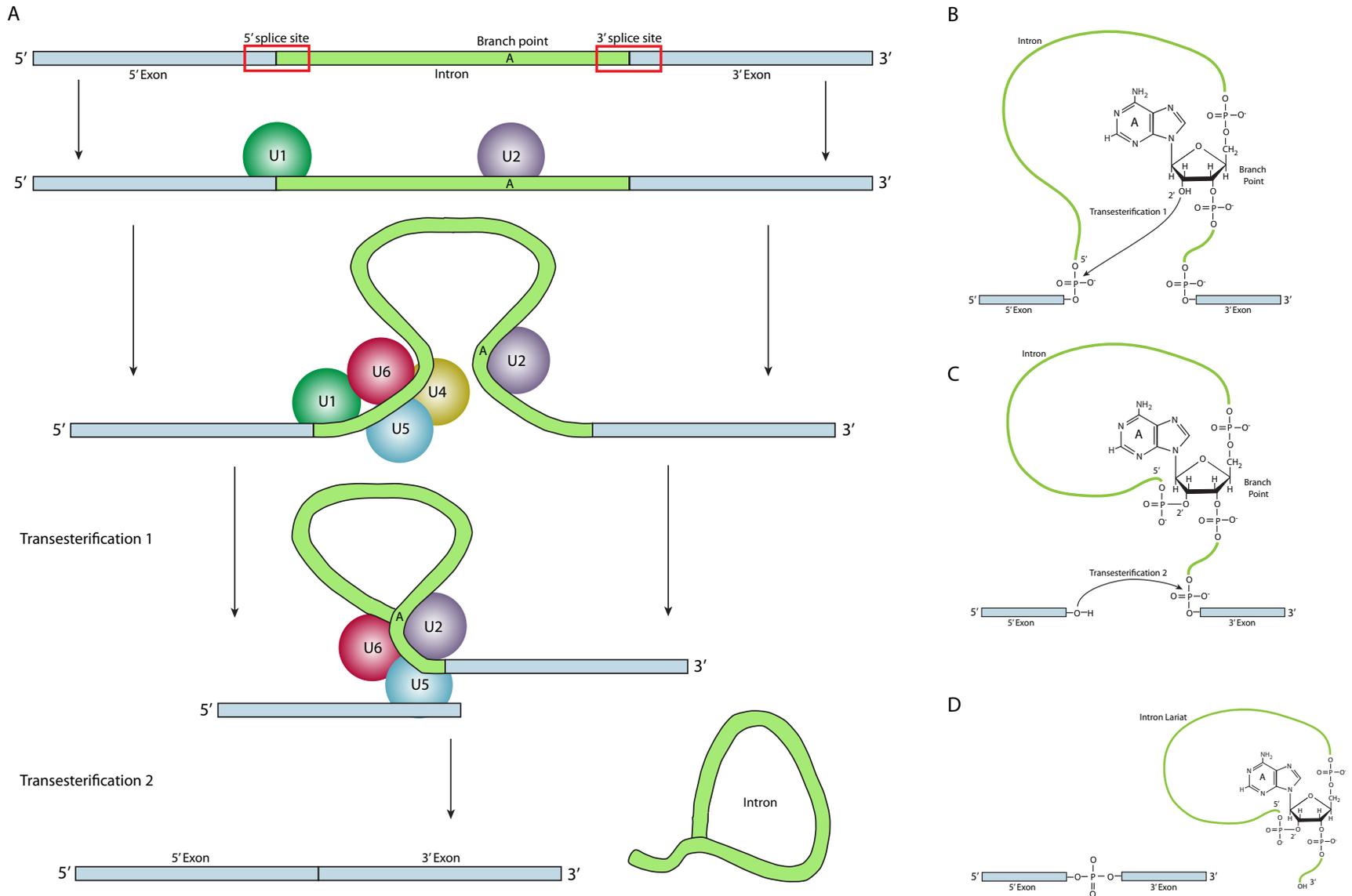


Figure 12. Splicing of RNA. Splicing removes some portions of a primary transcript (introns) while combining the remaining RNA (exons) to form the final mRNA sequence. The chemical mechanism is a series of two transesterifications. The first accomplishes the looping of the intron, while the second releases that intron as a “lariat” while simultaneously joining the two exons together. Splicing specificity relies on three landmarks on the RNA, most important of which is the branchpoint, an adenine residue that is the connection point for the intron loop formation. The other two landmarks are the 3' and the 5' splice sites, each of which are sequences that bind to snRNPs to bring together the spliceosome.

GENE REGULATION :

Control Mechanisms

To define a gene, a stretch of DNA must have a promoter, a start site, and a stop site. In a prokaryote, these are necessary and often sufficient, but in a eukaryote, they are still necessary, but seldom sufficient. This chapter discusses the other elements, both positive and negative, that are used to regulate the expression (i.e. transcription) of a gene. It is primarily a story of transcription factors and the recognition elements to which they bind.

Prokaryotic Transcriptional Regulation

Unlike multicellular organisms, in which most cells are in a tightly regulated internal environment, most prokaryotic cells are constantly responding to changing conditions in their immediate environment, such as changes in salt concentration, temperature, acidity, or nutrient availability. Because these organisms must respond quickly, the lifetime of an RNA is kept short, on the order of several minutes - so gene products that are not useful in the new conditions do not waste resources. For the same reason, initiation of new transcription must also occur very quickly - so that gene products that are needed to stabilize the cell in the new conditions are rapidly available. A fast and efficient control system is needed, and in prokaryotes, this means that the controls on transcription are simple activators and repressors. For some genes, both may be used for regulation, while for others, only one is needed to change from a default state of expression or non-expression.

A classic example of repressor control of gene expression, the lac operon, also illustrates another method by which bacteria may control the expression of genes. An *operon* is a group of genes whose products participate in the same metabolic pathway, and are transcribed under the control of a single promoter. The lac operon consists of three genes (lacZ, lacY, lacA) that participate in the catabolism of the disaccharide, lactose. LacZ is β -galactosidase, an enzyme that cleaves lactose into galactose and glucose. LacY is β -galactoside permease, which transports lactose from the extracellular environment into the cell. Both are required for lactose catabolism. Oddly, lacA is not absolutely

Using this book: This book is designed to be used in both introductory and advanced cell biology courses. The primary text is generally on the left side of the vertical divider, and printed in black. Details that are usually left to an advanced course are printed in blue and found on the right side of the divider. Finally, additional biomedically relevant information can be found in red print on either side of the divider.

required for lactose metabolism, but its function is related to the other two: it is a β -galactoside transacetylase that transfers acetyl groups from acetyl-CoA to lactose. All three are translated (they retain their individual start and stop codons for translation, not to be confused with the start and stop of transcription) from a single transcript. Of particular interest with respect to the regulation of this transcription is the structure of the promoter region. Note that in addition to the expected σ^{70} promoter upstream of the start site, there is another control sequence on each side of the start site (fig. 1A).

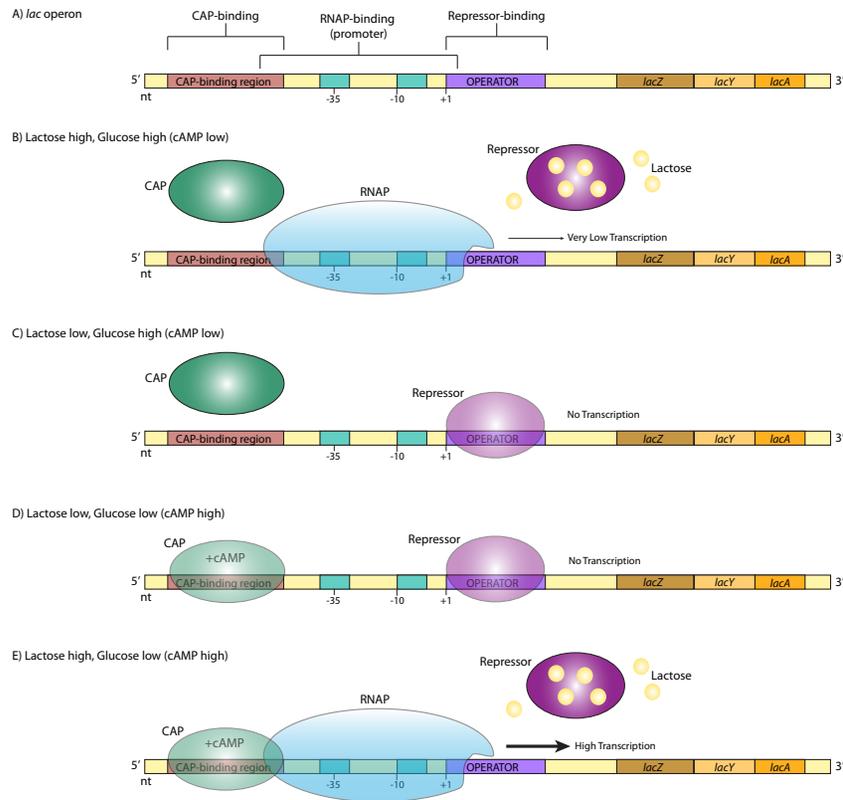


Figure 1. The *lac* operon.

The operator is a sequence of DNA that lies between the promoter and the start site. It is recognized by the lac repressor, a DNA binding protein with a helix-turn-helix motif. In the absence of lactose (fig. 1C), the lac repressor has a high affinity for the operator sequence and binds tightly, obstructing the start site and forming a physical “roadblock” to transcription by preventing the RNA polymerase from moving forward

Note that the helix-turn-helix (HTH) motif, which is common in bacterial DNA-binding proteins, is not the same thing as the helix-loop-helix DNA-binding proteins that are used in many eukaryotic systems. An elaboration of the basic HTH motif, known as the winged helix motif, is also found in a variety of prokaryotic DNA-binding proteins.

from the promoter. This makes sense physiologically because the cell is more efficient metabolizing glucose, and if there is no lactose around, then it is a waste of resources to make enzymes that metabolize it. However, what if there is suddenly an abundance of lactose in the environment? As the lactose is taken into the cell, intracellular levels rise, and now enzymes are needed to utilize this new food source. The lactose actually turns on the expression of enzymes that will metabolize it! Specifically, the lactose binds to the lac repressor protein (4 lactose binding sites), which causes a conformational change that releases it from the operator sequence (fig. 1B). Now an RNA polymerase that attaches at the lac operon promoter can proceed to transcribe the message unhindered, producing RNA and subsequently proteins that are used to break down the lactose. This continues as long as there is abundant lactose in the cell. As the lactose levels drop, repressor proteins are no longer bound by lactose, and can once again bind the operator and inhibit expression of the operon once again. For now, ignore the CAP protein in figure 1, and parts D and E. We'll come back to that. The lac operon is an example of an *inducible* operon, in which the native state is “off” and the introduction of and inducer (in this case lactose) will bind the repressor and turn the operon “on”.

In contrast, there are also operons with the reverse mechanism. An example of one such *repressible* operon is the *trp* operon (fig. 2). This operon contains five genes that are involved in the synthesis of the amino acid tryptophan: *trpE* and *trpD*, which

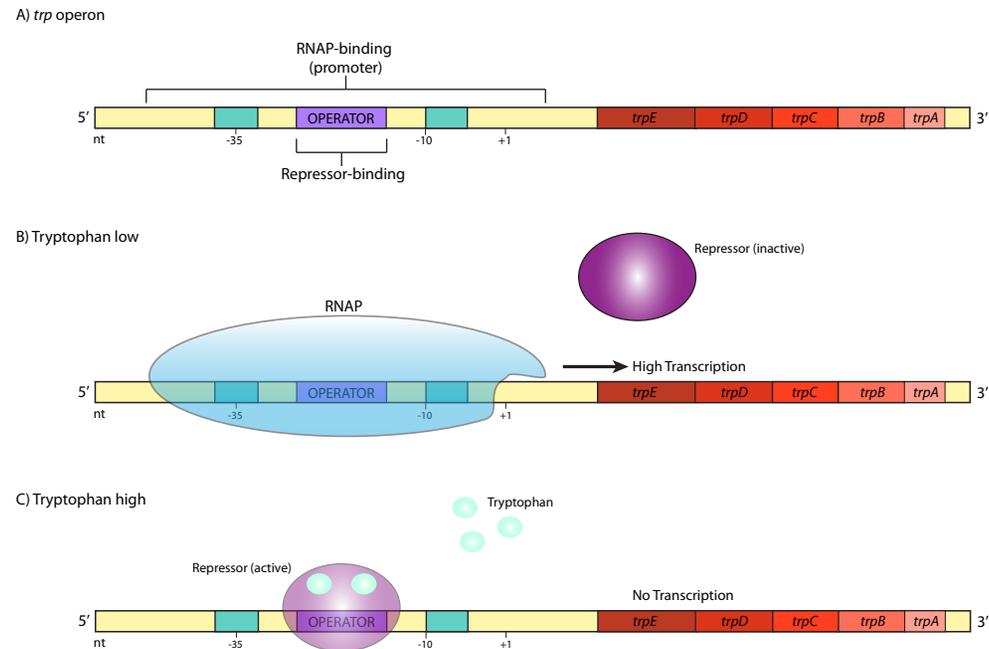


Figure 2. The *trp* operon.

together encode the subunits of anthranilate synthetase, *trpC*, which encodes N-(5'-phosphoribosyl)-anthranilate isomerase, and *trpB* and *trpA*, which each encode subunits of tryptophan synthetase. The *trp* repressor is larger and more complex than the *lac* repressor, but it also utilizes a helix-turn-helix DNA-binding motif.

However, it differs in a crucial aspect. In its native form, it does not bind to the operator sequence. It only binds to the operator after it has first bound tryptophan (two molecules of *trp* bind to one repressor). This is the opposite of the *lac* repressor, but when considering the physiological function of these genes, this should make perfect sense. As long as there is no tryptophan, the operator is unbound, allowing the RNA polymerase to transcribe the genes needed to make tryptophan (fig. 2B). When enough tryptophan has accumulated in the cell, some of the "extra" tryptophan binds to the *trp* repressor, which activates it and allows it to bind to the operator (fig. 2C). When this happens, the RNAP cannot reach the start site, and resources are not wasted transcribing genes for enzymes that make something the cell already has a lot of.

Let us now return to the *lac* operon in figure 1. It turns out that even when the operon is induced by the presence of lactose, the rate of transcription is low. The limitation is not from the repressor - that has been removed as described above (fig. 1B). Instead, the low expression is due to a low-affinity promoter. This is true not just of the *lac* operon, but also other non-glucose-pathway sugar-catabolism genes. There is a simple explanation: even if there are abundant alternate sugars available (e.g. lactose), if there is glucose available, it is the cell's most efficient and preferred pathway for energy production, and the production of enzymes for other pathways would be an inefficient use of resources. So, when and how is the *lac* operon really turned on?

The answer lies in a CAP, catabolite gene activator protein, also known as CRP, or cAMP receptor protein. It is a small homodimeric DNA binding protein that binds to a sequence that overlaps the 5' side of the promoter. In the presence of cAMP, which binds to the protein, CAP has a high affinity for the DNA recognition sequence, and binds to it (Fig. 1E). The protein then helps to recruit the RNAP to the promoter site, binding directly to the C-terminal domain of the RNAP α subunit to increase the affinity of the polymerase for the promoter sequence to overcome a weak promoter.

What does cAMP have to do with this? When there is abundant extracellular glucose, there is little cAMP. The enzyme that synthesizes cAMP, adenylate cyclase, is negatively regulated by glucose transport. However, when there is little environmental glucose, adenylate cyclase is more active, makes cAMP, which binds CAP, and leads to robust production of lactose catabolism enzymes. CAP is an example of an activator that can control gene expression in a positive direction.

In *E. coli*, cAMP levels are not directly tied to intracellular glucose levels or glucose metabolism. Rather, cAMP levels are altered by glucose transport through a phosphoenolpyruvate-dependent phosphotransferase system (PTS), part of which is de-phosphorylated (the *crr* gene product, also known as EIIA) when glucose is moved inward. The phosphorylated EIIA-P is an activator of adenylate cyclase. So, as glucose moves into the cell, cAMP levels drop due to inactive adenylate cyclase.

The last, and most complicated example of prokaryotic metabolic gene control is the *araBAD* operon. This operon produces enzymes used for the catabolism of the 5-carbon sugar, L-arabinose. The interesting thing about this operon is the presence of both positive and negative control elements that are used by the same control protein, *araC*. When there is little or no arabinose, the *araC* binds to the operator sequences *araO2* and *araI1*. The two *araC* proteins then interact, which causes the DNA to loop around preventing RNAP from binding to the promoter and transcribing *araBAD*. Furthermore, this operon is also under the control of CAP, and the double *araC* loop structure also

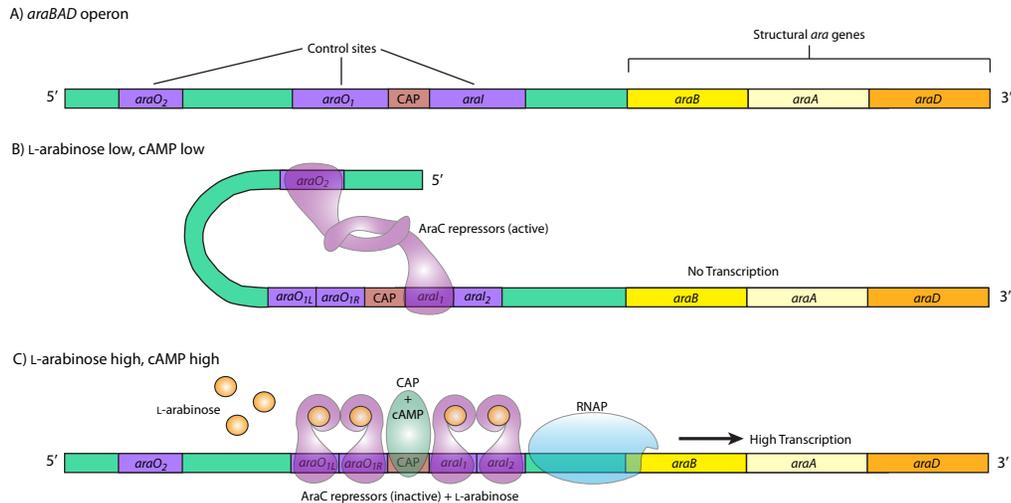


Figure 3. The *araBAD* operon.

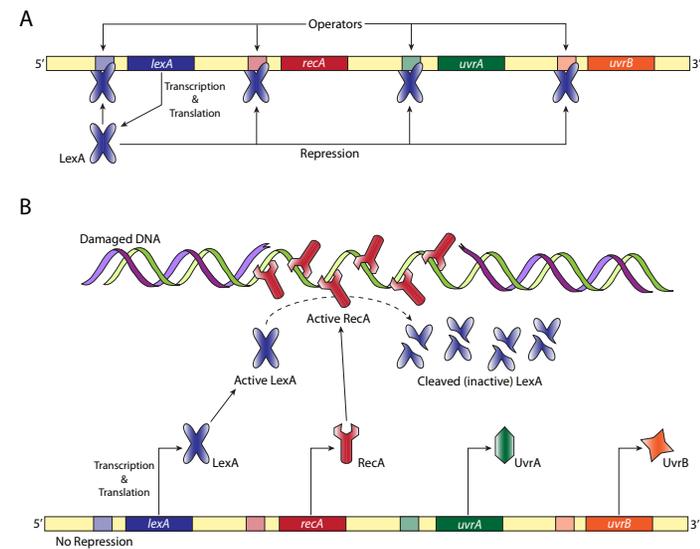
prevents CAP from binding. However, when there is plentiful arabinose, *araC* repressors bind the arabinose and then interact differently, still forming dimers, but now in a different conformation that leads to binding of *araO1L* and *araO1R* together as well as *araI1* and *araI2*. The arabinose-bound *araC* at the *araI* sites interact with RNAP and together with CAP promote strong activation of *araBAD* expression.

Eukaryotic Transcriptional Regulation

As with almost every comparison with prokaryotic systems, regulation of eukaryotic transcription is much more complex than prokaryotic gene control, although still based on similar mechanisms of activators and repressors. There is no close eukaryotic equivalent to operons, though: eukaryotic genes are always transcribed one per mRNA. The previous chapter described the formation of a preinitiation complex of transcription

Not all operons are concerned with coordinating metabolic activities. An important non-metabolic operon in *E. coli* is the *LexA/RecA* SOS response operon, which contains genes that are involved in DNA repair. The SOS repair system is invoked to allow DNA replication to continue through areas of damaged DNA, but with the penalty of low fidelity. One of the gene products of this operon, *RecA*, is important in recognizing and repairing damage caused by UV light. It also functions as a regulator of the *LexA* repressor protein. *LexA* is actually a repressor for multiple SOS operons, binding to a common operator sequence upstream of each gene/operon. It is activated when *RecA*, upon detecting DNA damage, undergoes a conformational shift and activates protease activity, which then cleaves *LexA*, allowing transcription from the SOS genes/operons.

SOS repair is error-prone because when the replisome encounters bulky damage, it undergoes “replication fork collapse” in which the DNA polymerase III units are released. The replacement, or bypass, polymerases, Pol IV (*dinB*), and Pol V (*umuDC*), do not have 3’–5’ proofreading exonuclease activity. Misincorporation of G opposite thymine dimers occurs at about half the rate of proper A incorporation, and generally, the bypass polymerases are about 1000 times more error-prone than Pol II or Pol I.



factors for RNA polymerase II. These transcription factors (e.g. TFIID, TFIIH, etc.) are known as general transcription factors, and are required for transcription of any gene at any level. However, there are also specific transcription factors, usually referred to simply as transcription factors (TF), that modulate the frequency of transcription of particular genes. Some upstream elements and their associated TFs are fairly common, while others are gene or gene-family specific. An example of the former is the upstream element AACCAAT and its associated transcription factor, CP1. Another transcription factor, Sp1, is similarly common, and binds to a consensus sequence of ACGCCC. Both are used in the control of the beta-globin gene, along with more specific transcription factors, such as GATA-1, which binds a consensus AAGTATCACT and is primarily produced in blood cells. This illustrates another option found in eukaryotic control that is not found in prokaryotes: tissue-specific gene expression. Genes, being in the DNA, are technically available to any and every cell, but obviously the needs of a blood cell differ a great deal from the needs of a liver cell, or a neuron. Therefore, each cell may produce transcription factors that are specific to its cell or tissue type. These transcription factors can then allow or repress expression of multiple genes that help define this particular cell type, assuming they all have the recognition sequences for the TFs. These recognition sequences are also known as response elements (RE).

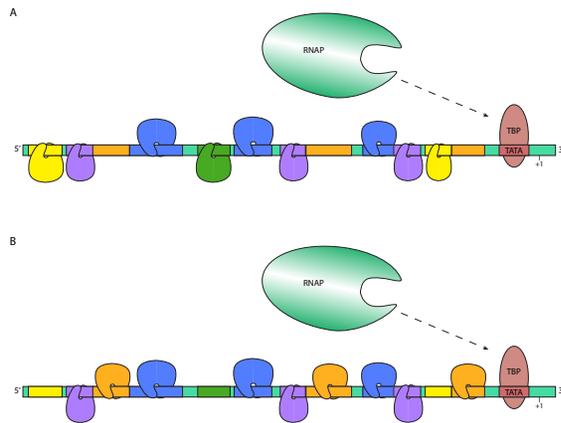


Figure 5. Eukaryotic transcription factors can work in complex combinations. In this figure, the transcription factors hanging downward are representative of inhibitory TFs, while those riding upright on the DNA are considered enhancers. Thus, the RNA polymerase in (A) has a lower probability of transcribing this gene, while the RNAP in (B) is more likely to, perhaps because the TF nearest the promoter interacts with the RNAP to stabilize its interactions with TFIID. In this way, the same gene may be expressed in very different amounts and at different times depending on the transcription factors expressed in a particular cell type.

Very often, a combination of many transcription factors, both enhancers and silencers, is responsible for the ultimate expression rate of a given eukaryotic gene. This can be done in a graded fashion, in which expression becomes stronger or weaker as more

enhancers or silencers are bound, respectively, or it can be a binary mode of control, in which a well-defined group of TFs are required to turn on transcription, and missing just one can effectively shut down transcription entirely. In the first case, activating TFs generally bind to the GTFs or RNA Polymerase II directly to help them recognize the promoter more efficiently or stably, while repressing TFs may bind to the activating TFs, or to the GTFs or RNAP II, in preventing recognition of the promoter, or destabilizing the RNAP II preinitiation complex. In the second case, activation hinges on the building of an enhanceosome, in which transcription factors and protein scaffolding elements and coactivators come together to position and stabilize the preinitiation complex and RNAP II on the promoter. The most prominent and nearly ubiquitous coactivator is named *Mediator*, and binds to the CTD of the β' subunit of RNA polymerase II and also to a variety of transcription factors.

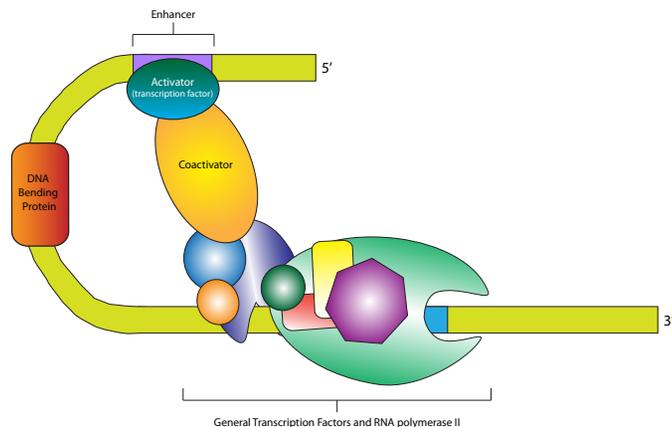


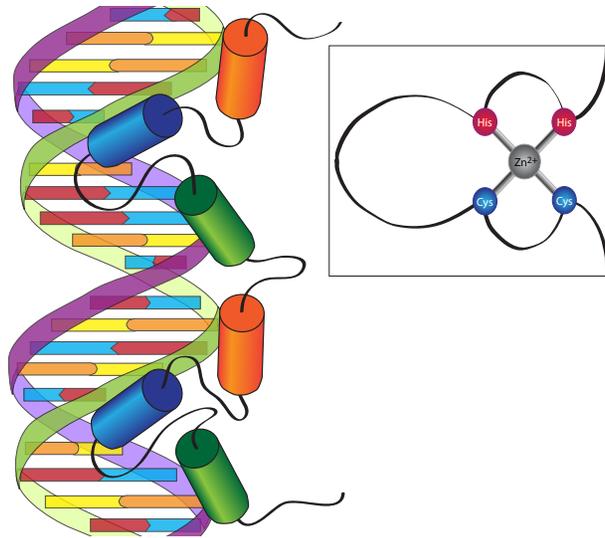
Figure 6. An enhancer can stabilize or recruit components of the transcription machine through a coactivator protein.

Eukaryotic transcription factors, while varied, usually contain at least one of the following transcription factor motifs: zinc fingers, leucine zippers, basic helix-loop-helix domains, Rel homology region domains, or a variation thereof.

The zinc finger motif was the first DNA-binding domain to be discovered, and was found in a general transcription factor associated with RNA polymerase III. The initial structure found was a repeating -30-amino acid motif with two invariant Cys and two invariant His residues that together bind a Zn^{++} ion and thus bring a tight loop or “finger” of basic potentially DNA-binding residues together. The basic finger binds to the major groove of the DNA, with the exact sequence-matching characteristics determined by the topology of the particular residues that make up the finger. Although most DNA binding motifs insert a positively-charged α -helical domain into the major groove of

DNA, the zinc-finger proteins are the only ones that combine several such motifs to interact with the DNA in several sequential sites.

Figure 7. Zinc-finger family transcription factor (left) and close-up of Zn^{2+} binding site (right). Each cylinder represents an α -helical domain.



The next motif is the leucine zipper. Although this is a common motif for transcription factors, it is important to note that unlike the zinc-finger, the leucine zipper itself is not a DNA-binding motif. Rather, it is a protein dimerization motif, and determines the way in which two protein subunits interact. However, the leucine zipper is a common structural motif in transcription factors. It works through opposing domains of regularly spaced hydrophobic amino acids, particularly leucines, which are very effective at holding the two subunits together in the aqueous environment of the cell. The leucines are found in every 7th residue position of an α -helical domain, leading to a coiled-coil superstructure when two subunits interact. The (+) charged DNA-binding domains of these proteins are usually N-terminal to the leucine zippers, as in the case of the bZIP category of leucine zipper proteins (the name stands for basic region leucine zipper).

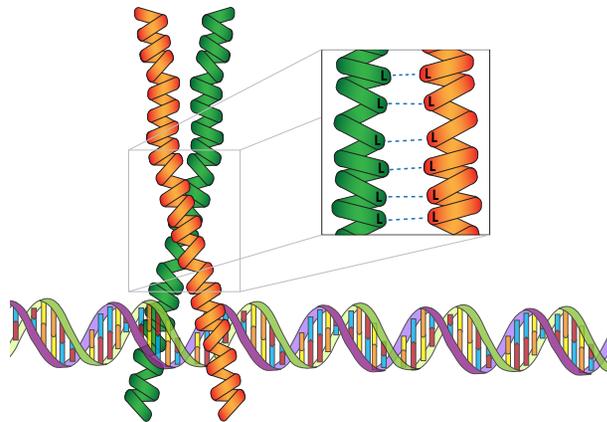


Figure 8. Leucine zipper.

In addition to the first type of Zn^{2+} -binding site described with two Cys and two His (Cys_2 - His_2), there are two major variations to note. The first is the Cys_2 - Cys_2 type, which is characteristic of steroid receptor transcription factors such as the glucocorticoid receptor or estrogen receptor. We will consider them in more detail later with the discussion of intracellular signal transduction, but for now, the general idea is that unactivated steroid hormone receptors are found in the cytoplasm, where they come in contact with and bind their cognate hormone molecule. They then translocate to the nucleus, where they dimerize and are able to act as transcription factors. The second major variation of the zinc finger is the binuclear Cys_6 , which carries six Cys residues to create a slightly larger "basket" in which two Zn^{2+} ions are held, rather than just one. The best-studied example of this type of zinc-finger protein is GAL4, a yeast metabolic transcription factor.

The bHLH, or basic helix-loop-helix domains appear to be elaborations on the leucine zipper theme. In this case, the N-terminal region is highly basic, making it ideal for interacting with DNA, and this basic domain, which is also helical, leads into the first helix (H1) of the motif, which is then connected by a non-helical loop of amino acids, leading into a second helical region (H2). Beyond the bHLH, these transcription factors may merge into a leucine zipper motif or other protein interaction domain for dimerization. Though the primary binding domain is N-terminal to H1, the H1 domains also appear to play a role in binding the major groove of the DNA. [Example myc]

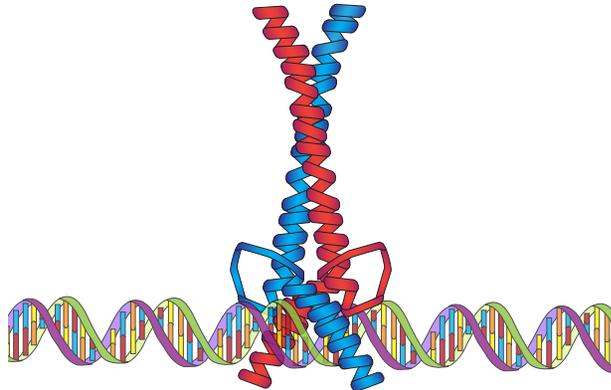


Figure 9. Binding of a basic helix-loop-helix (bHLH) class transcription factor.

NF- κ B (nuclear factor κ B) is a ubiquitous transcription factor discovered (and most noticeable) in the immune system. When active, it is a heterodimer, with both subunits containing a Rel homology region (RHR). Rel is an oncogene, and the RHR are named for their similarity to the previously-sequenced rel. The RHR domains bind to DNA with extraordinary affinity, due in part to having five loops for DNA contact per subunit. Just as with the other types of transcription factors, some RHR-containing proteins are repressors, while others are activators.

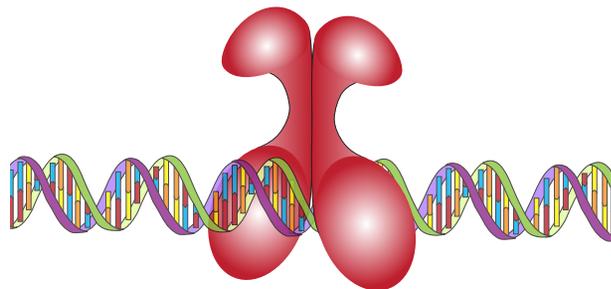


Figure 10. RHR domains are a DNA-binding domain found in the NF- κ B family of transcription factors.

The regulation of NF- κ B is rather interesting: once it is in the nucleus, it is generally active. However, it is, as almost all cellular proteins, made in the cytoplasm. Inhibitors of NF- κ B (I κ B) also reside in the cytoplasm, and they act by binding the NF- κ B and covering the nuclear localization signal that allows its import into the nucleus. Thus sequestered, the NF- κ B must remain in the cytoplasm inactive until some stimulus activates I κ B kinase, which phosphorylates the I κ B and leads to ubiquitination and degradation, finally releasing the NF- κ B from its bonds.

Because it can be mobilized quickly (compared to synthesizing new protein), NF- κ B is considered a rapid-response transcription factor that is often used to begin expression of a gene needed soon after it has been “ordered” by a signal, either extracellular or intracellular. Not surprisingly for a factor discovered in the immune system, it is activated in response to bacterial and viral antigens, as well as other types of cellular stress or insult.

In addition to the relatively short-term regulation of gene expression controlled by binding transcription factors to regulatory elements, there are also stronger methods of locking away a gene to prevent its expression. In chapter 7, acetylation and deacetylation of histones was discussed as a method for decreasing and increasing their affinity for DNA. This can be controlled (Fig. 11B) by the recruitment of histone deacetylase (HDAC) to particular genes via repressor/co-repressor complexes. The deacetylase forces tight winding of the targeted DNA to the histones, precluding access by RNA polymerases or general transcription factors.

Another recruiter of HDAC are MBD proteins, which bind to methylated DNA. DNA methylation in mammals usually occurs on CpG dinucleotide sequences. This methylation appears to have the effect of blocking access of transcription factors and enzymes to the DNA. It can do so directly, or by recruiting MBD (methyl-CpG-binding domain) proteins. In either case, methylation is a long-term method of locking up genes, and is the mechanism for turning off genes that would never be used in a particular cell type (e.g. hemoglobin in neurons).

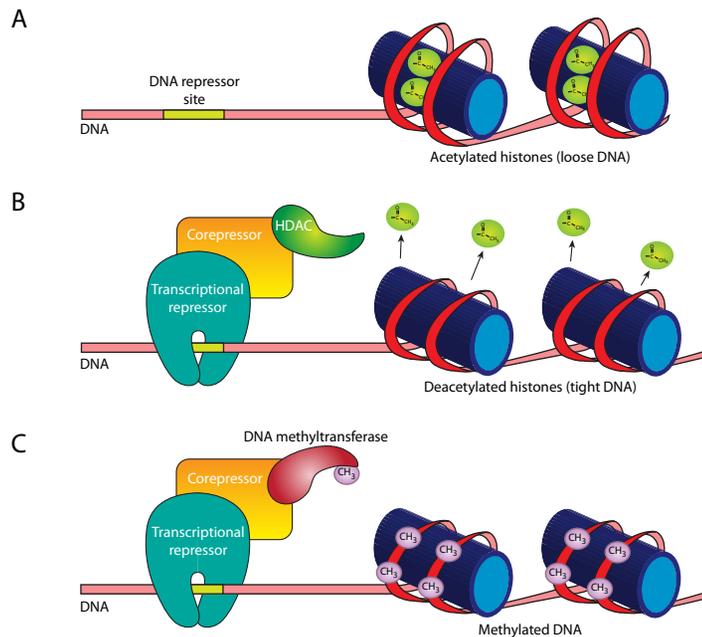


Figure 11. Long-term gene repression. (A) Acetylation of histones allows DNA to move off, potentially freeing the genes in that region for expression. (B) Specific regions can be wrapped more tightly around histones through the action of HDAC, removing the acetyl groups. (C) Methylation of the DNA can also prevent expression, either by physically blocking access, or by recruiting HDAC.

TRANSLATION :

From RNA to Protein

The RNA polymerase has done its job (or in the case of prokaryotes, may still be in the process of doing its job), so now what happens to the RNA? For RNA that is destined to provide instructions for making a protein, then it needs to be translated, which is a job for Superman™! Oops, actually it's a job for ribosomes.

Ribosomes are a complex of RNA and protein that bind to and processively move down (from 5' to 3' end) a strand of mRNA, picking up aminoacyl-tRNAs, checking to see if they are complementary to the RNA tri-nucleotide being “read” at the moment, and adding them to the new polypeptide chain if they are. The RNA part of the ribosomes are generated by the organism's general purpose RNA polymerase in prokaryotes, and generated by the RNA polymerases I and III in eukaryotes. Recall that RNA Pol II is used by eukaryotes to generate protein-coding mRNA's. Although the numbers of RNA strands and protein subunits differ between the prokaryote and eukaryote, the mechanism for translation is remarkably well conserved.

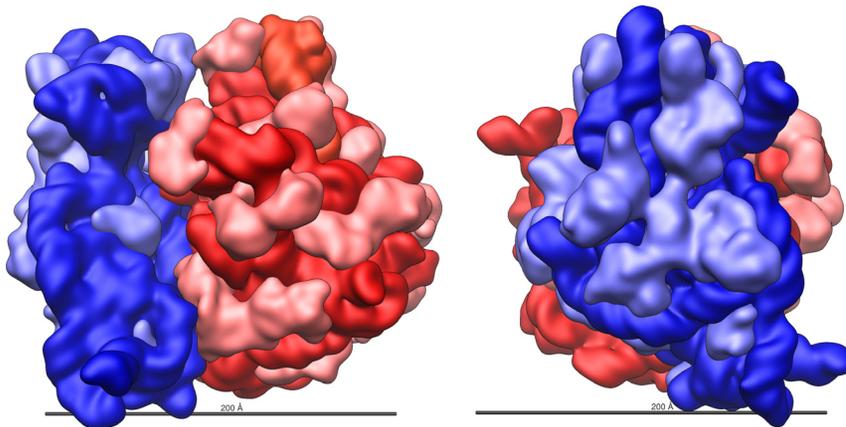


Figure 1. Two views of a prokaryotic ribosome. The large ribosomal subunit (50S) is shown in red, while the small ribosomal subunit (30S) is shown in blue. 3D images generated from data in the RCSB Protein Data Bank.

Using this book: This book is designed to be used in both introductory and advanced cell biology courses. The primary text is generally on the left side of the vertical divider, and printed in black. Details that are usually left to an advanced course are printed in blue and found on the right side of the divider. Finally, additional biomedically relevant information can be found in red print on either side of the divider.

Prokaryotic ribosomes

The prokaryotic ribosomes contain 3 RNA strands and 52 protein subunits which can be divided into 1 RNA and 21 proteins in the small ribosomal subunit (aka the 30S subunit) and 2 RNA and 31 proteins in the large ribosomal subunit (50S subunit). The small subunit locates the start site and moves along the RNA. The large ribosomal subunit contains the aminoacyl transferase enzyme activity that connects amino acids to make a protein. Neither subunit is sufficient to carry out translation by itself. They must come together to form the full 70S ribosome for translation to occur. If you aren't already familiar with the nomenclature, you're probably thinking that it's obvious why I went into biology rather than math. My competence at basic computation aside, there is a method to the madness. The "S" in 30S or 50S indicates Svedberg units, or a measurement of the sedimentation rate when the molecules in question are centrifuged under standard conditions. Because the rate of sedimentation depends on both the mass *and the shape* of a molecule, numbers do not always add up.

The genes for the prokaryotic rRNA molecules are arranged in an operon and thus come from a single transcript. Depending on the organism, there may be several such operons in the genome to ensure steady production of this crucial enzymatic complex. However, these RNAs are not translated, so instead of having multiple translation start codons to signal the beginning of each gene, the single transcript is cleaved post-transcriptionally by Ribonuclease III (RNase III) into 23S, 16S, and 5S segments, and these are then further trimmed by RNase III and RNase M into the final 23S, 16S, and 5S rRNAs found in the ribosomes. Although the 5S does not appreciably differ in sedimentation rate, it is in fact slightly "shaved" in post-transcriptional editing.

Eukaryotic ribosomes

Like the RNA molecules in prokaryotic ribosomes, the eukaryotic rRNA molecules are also post-transcriptionally cleaved from larger transcripts. This processing, and the subsequent assembly of the large and small ribosomal subunits are carried out in the nucleolus, a region of the nucleus specialized for ribosome production, and containing not only high concentrations of rRNA and ribosomal proteins, but also RNA polymerase I and RNA polymerase III. In contrast, RNA polymerase II, as befits its broader purpose, is found throughout the nucleus. The 40S small ribosomal subunit in eukaryotes also has just 1 rRNA, and has 33 proteins. The 60S, or large ribosomal subunit in eukaryotes has three rRNA molecules, two of which are roughly analogous to the prokaryote (28S and 5S eukaryotic, 23S and 5S prokaryotic), and one, the 5.8S, that binds with complemen-

Because of the density of material in the nucleolus needed for constant ribosome production, it is often readily visible under various types of microscopy despite not being bounded by a membrane.

tary sequence on part of the 28S rRNA. It also contains 50 proteins. These ribosomal subunits have roughly the same function as the prokaryotic versions: the small subunit in conjunction with various initiation factors is responsible for finding the start site and positioning the ribosome on the mRNA, while the large subunit houses the docking sites for incoming and spent aminoacyl-tRNAs and contains the catalytic component to attach amino acids via peptide bonds.

The ribosomal RNA precursors (pre-rRNA) are remarkably conserved in eukaryotes, with the 28S, 5.8S, and 18S rRNAs encoded within a single transcript. This transcript, synthesized by RNA polymerase I, always has the same 5' to 3' order: 18S, 5.8S, 28S. After the 45S pre-rRNA is transcribed, it is immediately bound by nucleolar proteins in preparation for cleavage and base modification. However, it is primarily the small nucleolar RNAs (snoRNAs), not the proteins, that determine the position of the modifications. The primary modifications are 2'-hydroxymethylation and transformation of some uridines into pseudouridines. These snoRNAs are sometimes transcribed independently by RNA polymerase III or II, but often are formed from the introns of pre-mRNA transcripts. Oddly, some of these introns come from pre-mRNAs that form unused mRNAs!

For the nucleolus to be the site of ribosome assembly, the ribosomal proteins must be available to interact with the rRNAs. By mechanisms discussed in the next chapter, the mRNAs for the ribosomal proteins are translated (as are all proteins) in the cytoplasm, but the resulting proteins are then imported into the nucleus for assembly into either the large or small ribosomal subunit. The subunits are then exported back out to the cytoplasm, where they can carry out their function.

The Genetic Code

We have blithely described the purpose of the DNA chromosomes as carrying the information for building the proteins of the cell, and the RNA as the intermediary for doing so. Exactly how is it, though, that a molecule made up of just four different nucleotides joined together (albeit thousands and even thousands of thousands of them), can tell the cell which of twenty-odd amino acids to string together to form a functional protein? The obvious solution was that since there are not enough individual unique nucleotides to code for each amino acid, there must be combinations of nucleotides that designate particular amino acids. A doublet code, would allow for only 16 different combinations (4 possible nucleotides in the first position x 4 possible nucleotides in the 2nd position = 16 combinations) and would not be enough to encode the 20 amino acids. However, a triplet code would yield 64 combinations, easily enough to encode

20 amino acids. So would a quadruplet or quintuplet code, for that matter, but those would be wasteful of resources, and thus less likely. Further investigation proved the existence of a triplet code as described in the table below.

		1st base								
		U		C		A		G		
2nd base	U	UUU	Phenylalanine	UCU	Serine	UAU	Tyrosine	UGU	Cysteine	U
		UUC	Phenylalanine	UCC	Serine	UAC	Tyrosine	UGC	Cysteine	C
		UUA	Leucine	UCA	Serine	UAA	Stop	UGA	Stop	A
		UUG	Leucine	UCG	Serine	UAG	Stop	UGG	Tryptophan	G
C	CUU	Leucine	CCU	Proline	CAU	Histidine	CGU	Arginine	U	
	CUC	Leucine	CCC	Proline	CAC	Histidine	CGC	Arginine	C	
	CUA	Leucine	CCA	Proline	CAA	Glutamine	CGA	Arginine	A	
	CUG	Leucine	CCG	Proline	CAG	Glutamine	CGG	Arginine	G	
A	AUU	Isoleucine	ACU	Threonine	AAU	Asparagine	AGU	Serine	U	
	AUC	Isoleucine	ACC	Threonine	AAC	Asparagine	AGC	Serine	C	
	AUA	Isoleucine	ACA	Threonine	AAA	Lysine	AGA	Arginine	A	
	AUG	Methionine (Start)	ACG	Threonine	AAG	Lysine	AGG	Arginine	G	
G	GUU	Valine	GCU	Alanine	GAU	Aspartic Acid	GGU	Glycine	U	
	GUC	Valine	GCC	Alanine	GAC	Aspartic Acid	GGC	Glycine	C	
	GUA	Valine	GCA	Alanine	GAA	Glutamic Acid	GGA	Glycine	A	
	GUG	Valine	GCG	Alanine	GAG	Glutamic Acid	GGG	Glycine	G	

Nonpolar, aliphatic
 Polar, uncharged
 Aromatic
 Positively charged
 Negatively charged

With so many combinations and only 20 amino acids, what does the cell do with the other possibilities? The genetic code is a *degenerate code*, which means that there is redundancy so that most amino acids are encoded by more than one triplet combination (codon). Although it is a redundant code, it is not an ambiguous code: under normal circumstances, a given codon encodes one and only one amino acid. In addition to the 20 amino acids, there are also three “stop codons” dedicated to ending translation. The three stop codons also have colloquial names: UAA (ochre), UAG (amber), UGA (opal), with UAA being the most common in prokaryotic genes. Note that there are no dedicated start codons: instead, AUG codes for both methionine and the start of translation, depending on the circumstance, as explained forthwith. The initial Met is a methionine, but in prokaryotes, it is a specially modified formyl-methionine (f-Met). The tRNA is also specialized and is different from the tRNA that carries methionine to the ribosome for addition to a growing polypeptide. Therefore, in referring to a the loaded initiator tRNA, the usual nomenclature is fMet-tRNA_i or fMet-tRNA_f. There also seems to be a little more leeway in defining the start site in prokaryotes than in eukaryotes, as some bacteria use GUG or UUG. Though these codons normally encode valine and leucine, respectively, when they are used as start codons, the initiator tRNA brings in f-Met.

The colloquial names were started when the discoverers of UAG decided to name the codon after a friend whose last name translated into “amber”. Opal and ochre were named to continue the idea of giving stop codons color names.

The stop codons are sometimes also used to encode what are now considered the 21st and 22nd amino acids, selenocysteine (UGA) and pyrrolysine (UAG). These amino acids have been discovered to be consistently encoded in some species of prokarya and archaea.

Although the genetic code as described is nearly universal, there are some situations in which it has been modified, and the modifications retained in evolutionarily stable environments. The mitochondria in a broad range of organisms demonstrate stable changes to the genetic code including converting the AGA from encoding arginine into a stop codon and changing AAA from encoding lysine to encoding asparagine. Rarely, a change is found in translation of an organismic (nuclear) genome, but most of those rare alterations are conversions to or from stop codons.

tRNAs are rather odd ducks

In prokaryotes, tRNA can be found either as single genes or as parts of operons that can also contain combinations of mRNAs or rRNAs. In any case, whether from a single gene, or after the initial cleavage to separate the tRNA transcript from the rest of the transcript, the resulting pre-tRNA has an N-terminal leader (41 nt in *E. coli*) that is excised by RNase P. That cleavage is universal for any prokaryotic tRNA. After that, there are variations in the minor excisions carried out by a variety of nucleases that produce the tRNA in its final length though not its mature sequence, as we will see in a few paragraphs.

Eukaryotic pre-tRNA (transcribed by RNA polymerase III) similarly has an N-terminal leader removed by RNase P. Unlike the prokaryote though, the length can vary between different tRNAs of the same species. Some eukaryotic pre-tRNA transcripts also contain introns, especially in the anticodon loop, that must be spliced out for the tRNA to function normally. These introns are different from the self-splicing or spliceosome-spliced transcripts discussed in the transcription chapter. Here, the splicing function is carried out not by ribozymes, but by conventional (protein) enzymes. Interestingly, RNaseP also removes a 3' sequence from the pre-tRNA, but then another 3' sequence is added back on. This new 3' end is always CCA, and is added by three successive rounds with tRNA nucleotidyl transferase.

Earlier in this textbook, when RNA was introduced, it was noted that although extremely similar to DNA on many counts, it is normally single stranded, and that property, combined with the opportunity for complementary base pairing within a strand, allows it to do something far different than double-stranded DNA: it can form highly complex secondary structures. One of the simplest and clearest examples of this is tRNA, which depends on its conformation to accomplish its cellular function. The prototypical cloverleaf-form tRNA diagram is shown in fig. 2 on the left, with a 3D model derived from x-ray crystallographic data on the right. As you can see, the fully-splayed-out shape

Other minor alterations to the genetic code exist as well, but the universality of the code in general remains. Some mitochondrial DNAs can use different start codons: human mitochondrial ribosomes can use AUA and AUU. In some yeast species, the CGA and CGC codons for arginine are unused. Many of these changes have been cataloged by the National Center for Biotechnology Information (NCBI) based on work by Jukes and Osawa at the University of California at Berkeley (USA) and the University of Nagoya (Japan), respectively.

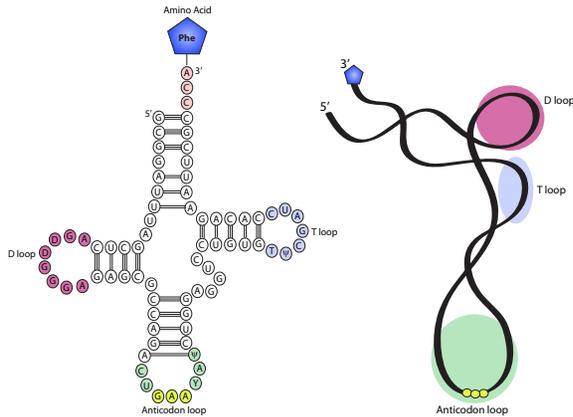


Figure 2. tRNA structure. At left, the classic clover-leaf, splayed out for simplicity. At right, a more accurate representation of the tRNA in pseudo-3D.

has four stem-loop “arms” with the amino acid attached to the acceptor arm, which is on the opposite side of the tRNA from the anticodon arm, which is where the tRNA must match with the mRNA codon during translation. Roughly perpendicular to the acceptor-anticodon axis are the D arm and the T ψ C arm. In some tRNAs, there are actually five total arms with a very short loop between the T ψ C and anticodon arms. The arm-like stem and loop structures are formed by two areas of strong complementarity (the stems, base-paired together) interrupted by a short non-complementary sequence (the loop). In general terms, the arms are used to properly position the tRNA within the ribosome as well as recognizing the mRNA codon and bringing in the correct amino acid.

When it comes time for the tRNA to match its anticodon with the codon on the mRNA, the code is not followed “to the letter” if you will pardon the pun. There is a phenomenon called “wobble” in which a codon-anticodon match is allowed and stabilized for translation even if the nucleotide in the third position is not complementary. Wobble can occur because the conformation of the tRNA allows a little flexibility to that position of the anticodon, permitting H-bonds to form where they normally would not. This is not a universal phenomenon though: it only applies to situations where a U or a G is in the first position of the anticodon (matching the third position of codon). Following the convention of nucleic acid sequences, the sequence is always written 5’ to 3’, even though in the case of codon-anticodon matching, the strands of mRNA and tRNA are antiparallel:

```

tRNA  3' U A A 5'
mRNA  5' A U U 3'

```

In addition to being allowed a bit of wobble in complementary base-pairing, tRNA molecules have another peculiarity. After being initially incorporated into a tRNA through conventional transcription, there is extensive modification of some of the bases of the tRNA. This affects both purines and pyrimidines, and can range from simple additions such as methylation or extensive restructuring of the sugar skeleton itself, as in the conversion of guanosine to wyosine (W). Over 50 different modifications have been catalogued to date. These modifications can be nearly universal, such as the dihydrouridine (D) found in the tRNA D loop, or more specific, such as the G to W conversion found primarily in tRNA^{Phe} of certain species (examples have been identified in both prokaryotic and eukaryotic species). As many as 10% of the bases in a tRNA may be modified. Naturally, alterations to the sugar base of the nucleotides can also alter the base-pairing characteristics. For example, one common modified base, inosine, can complement U, C, or A. This aberrant complementary base pairing can be equal among the suitor bases, or it may be biased, as in the case of 5-methoxyuridine, which can recognize A, G, or U, but the recognition of U is poor.

The knowledge of the genetic code begs the question: how is the correct amino acid attached to any given tRNA? A class of enzymes called the *aminoacyl tRNA synthetases* are responsible for recognizing both a specific tRNA and a specific amino acid, binding an ATP for energy and then joining them together (sometimes called charging the tRNA) with hydrolysis of the ATP. Specificity is a difficult task for the synthetase since amino acids are built from the same backbone and are so similar in mass. Distinguishing between tRNA molecules is easier, since they are larger and their secondary structures also allow for greater variation and therefore greater ease of discrimination. There is also a built-in pre-attachment proofreading mechanism in that tRNA molecules that fit the synthetase well (i.e. the correct ones) maintain contact longer and allow the reaction to proceed whereas ill-fitting and incorrect tRNA molecules are likely to disassociate from the synthetase before it tries to attach the amino acid.

Charging an aminoacyl tRNA synthetase with its amino acid requires energy. The synthetase first binds a molecule of ATP and the appropriate amino acid, which react resulting in the formation of aminoacyl-adenylate and pyrophosphate. The PPi is released and the synthetase now binds to the proper tRNA. Finally, the amino acid is transferred to the tRNA. Depending on the class of synthetase, the amino acid may attach to the 2'-OH of the terminal A (class I) or to the 3'-OH of the terminal A (class II) of the tRNA. Phe-tRNA synthetase is the exception: it is structurally a class II enzyme but transfers the Phe onto the 2'-OH. Note that amino acids transferred onto the 2'-OH are soon moved to the 3'-OH anyway due to a transesterification reaction.

Prokaryotic Translation

As soon as the RNA has emerged from the RNAP and there is sufficient space to accommodate a ribosome, translation can begin in prokaryotes. In fact, for highly expressed genes, it would not be unusual to see multiple RNA polymerases transcribing the DNA and multiple ribosomes on each of the transcripts translating the mRNA to protein! The process begins with the small ribosomal subunit (and only the small subunit - if it is attached to the large subunit, it is unable to bind the mRNA), which binds to the mRNA loosely and starts to scan it for a recognition sequence called the Shine-Dalgarno sequence, after its discoverers. Once this is recognized by the small ribosomal subunit rRNA, the small subunit is positioned around the start codon (AUG). This process is facilitated by initiation factors as follows. The 30S ribosomal subunit dissociates from the 50S ribosomal subunit if it was associated with one, and binds to initiation factors IF-1 and IF-3. IF-1 binds to the A site, where it prevents new aminoacyl-tRNA molecules from entering before the full ribosome is assembled. It also facilitates the assembly and stabilization of the initiation complex. IF-3 is required to allow the 30S subunit to bind to mRNA. Once this has occurred, IF-2-GTP arrives on scene, carrying with it the initiator aminoacyl-tRNA. This settles into the P site, which is positioned so that the anticodon of the tRNA settles over the AUG start codon of the mRNA. Hydrolysis of the GTP attached to IF-2 and release of all the initiation factors is needed to allow the 50S subunit to bind to the 30S subunit to form the full and fully functional ribosome. Because GTP hydrolysis was required, the joining of the subunits is irreversible spontaneously, and requires expenditure of energy upon termination of translation. Once the 50S subunit joins with the 30S subunit, the A site is ready to accept the next aminoacyl-tRNA.

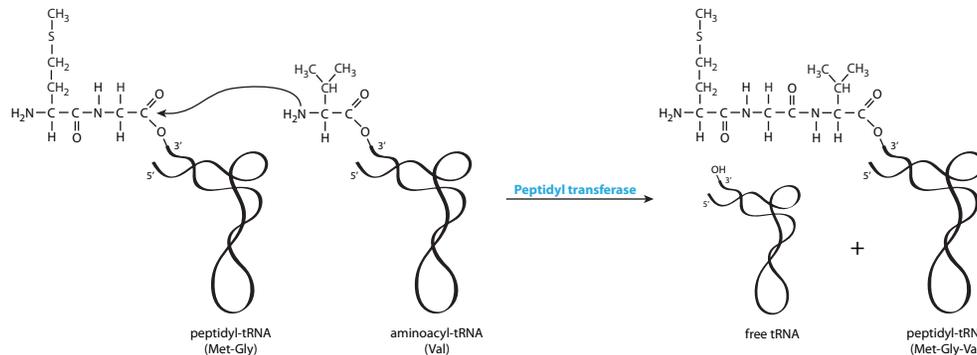


Figure 4. Peptide bond formation at the addition of the third amino acid. The previous two amino acids are peptide bonded together as well as attached to the tRNA from the second amino acid. The aminoacyl-tRNA bond is broken and transferred/transformed to the peptide bond connecting the initial dipeptide to the third amino acid.

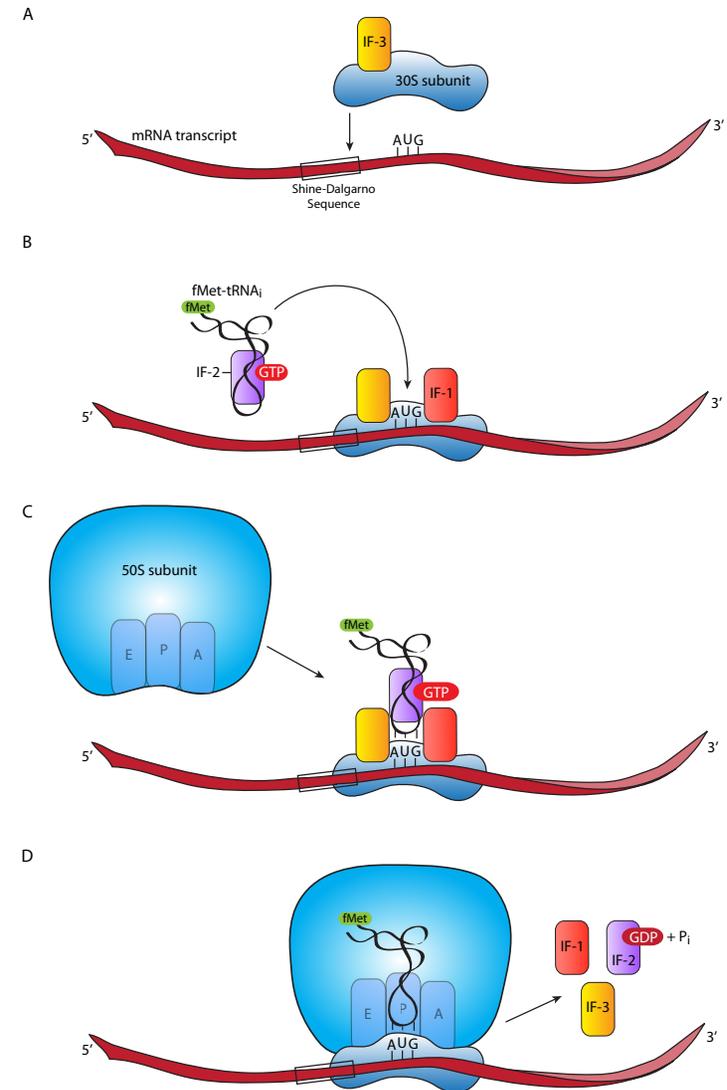


Figure 3. Initiation of Translation in Prokaryotes. (A) 30S subunit binds to Shine-Dalgarno sequence. (B) fMet-tRNA_i is loaded into the middle slot of the small ribosomal subunit. Initiation factors occupy the other two slots. (C) The large ribosomal subunit docks with the small subunit. (D) The initiation factors are released and the ribosome is ready to start translation.

A common and understandable misconception is that the new amino acid brought to the ribosome is added onto the growing polypeptide chain. In fact, the mechanism is exactly the opposite: the polypeptide is added onto the new amino acid (fig. 4). This begins from the second amino acid to be added to a new protein (fig. 5). The first amino acid, a methionine, you should recall, came in along with IF-2 and the initiator tRNA. The new aminoacyl-tRNA is escorted by EF-Tu, an elongation factor that carries a GTP. Once the aa-tRNA is in place, EF-Tu hydrolyzes the GTP and dissociates from the aminoacyl-tRNA and ribosome.

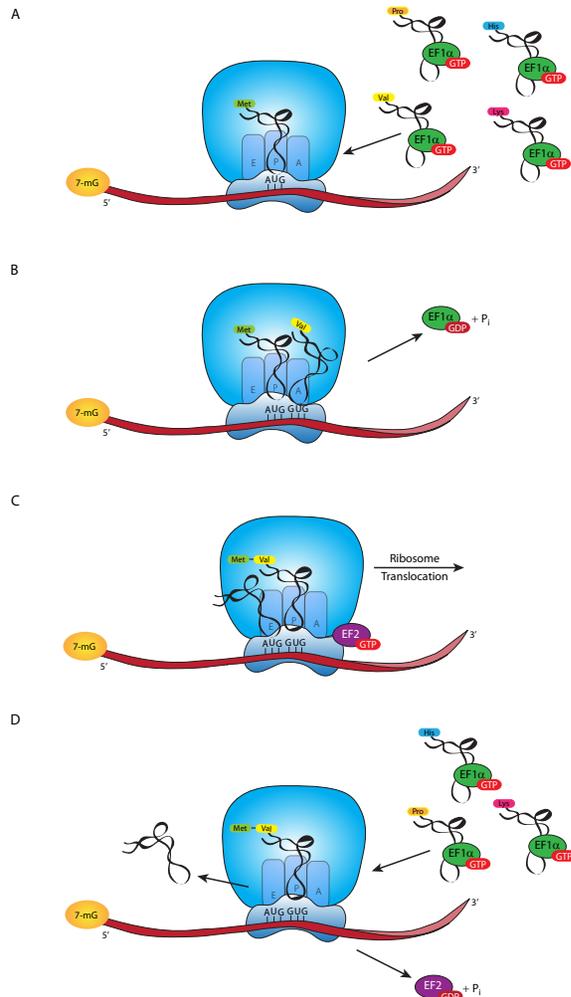


Figure 5. Elongation of the polypeptide chain. (A) a new aminoacyl-tRNA drops into the A slot of the ribosome. (B) the fMet is moved off its tRNA_i and peptide-bonded to the new amino acid, which is still attached to its tRNA. (C) Ribosome shifts over to the right. (D) empty tRNA_i is ejected.

For a long time, there was a bit of mystery surrounding the simultaneous docking of two tRNA molecules on immediately adjacent codons of mRNA. Under normal conditions, there should not be enough room, since the tRNAs are fairly bulky and one should obstruct the other from reaching the mRNA to make a codon-anticodon match. The matter was finally cleared up in 2001 with x-ray crystallographic examinations showing a bend in the mRNA between the codon in the P slot and the codon in the A slot. The bend puts the two associated tRNAs at slightly different angles and thus creates just enough room for both to maintain basepairing hydrogen bonds with the mRNA. See Yusupov et al, *Science* 292 (5518): 883-896, 2001.

When a new aminoacyl-tRNA drops into the A slot of the ribosome, the anticodon is lined up with the codon of the mRNA. If there is no complementarity, the aminoacyl-tRNA soon floats back out of the slot to be replaced by another candidate. However, if there is complementarity (or something close enough, recalling the idea of wobble) then H-bonds form between the codon and anti-codon, the tRNA changes conformation, which shifts the conformation of EF-Tu, causing hydrolysis of GTP to GDP + Pi, and release from the aa-tRNA. The codon-anticodon interaction is stable long enough for the catalytic activity of the ribosome to hydrolyze the bond between fMet and the tRNAf in the P slot, and attach the fMet to the new amino acid with a peptide bond in the A slot. The new amino acid is still attached to its tRNA, and as this process occurs, the ribosome shifts position with respect to the mRNA and tRNAs. This puts the now-empty (no amino acid attached) tRNAf in the E slot, the tRNAaa in the P slot, attached to that aa which is bonded to Met, and the A slot is again open for a new tRNA to come in. The elongation factor EF-G binds near the A slot as soon as EF-Tu leaves, and is required for ribosomal translocation, providing energy for the process by hydrolyzing a GTP that it carries with it to the ribosome. From my students' experiences, the best way to learn this seems to be to study the diagrams and see the movements of the molecules, filling in the mechanistic details in your mind. This process continues until the ribosome brings the A slot in line with a stop codon.

There is no tRNA with an anticodon for the stop codon. Instead, there is a set of release factors that fit into the A site of the ribosome, bind to the stop codon, and activate the ribosome to cut the bond between the polypeptide chain and the last tRNA (fig. 6). Depending on which stop codon is present either RF1 (recognizes UAA or UAG) or RF2 (for UAA or UGA) first enters the A slot. The RF1 or RF2 is complexed with RF3, which is involved in subsequent releasing of the RF complex from the A slot. This is necessary because once the polypeptide has been released from the ribosome, the mRNA must be released. Ribosome releasing factor (RRF) also binds in the A slot, which causes a conformational change in the ribosome releasing the previous and now empty tRNA. Finally, EF-G binds to RRF, and with an accompanying hydrolysis of GTP, causes dissociation of the ribosome into separate large and small subunits. Note that it is the combination of EF-G/RRF that causes dissociation; EF-G alone plays a different role in ribosome movement when it is not at the stop codon.

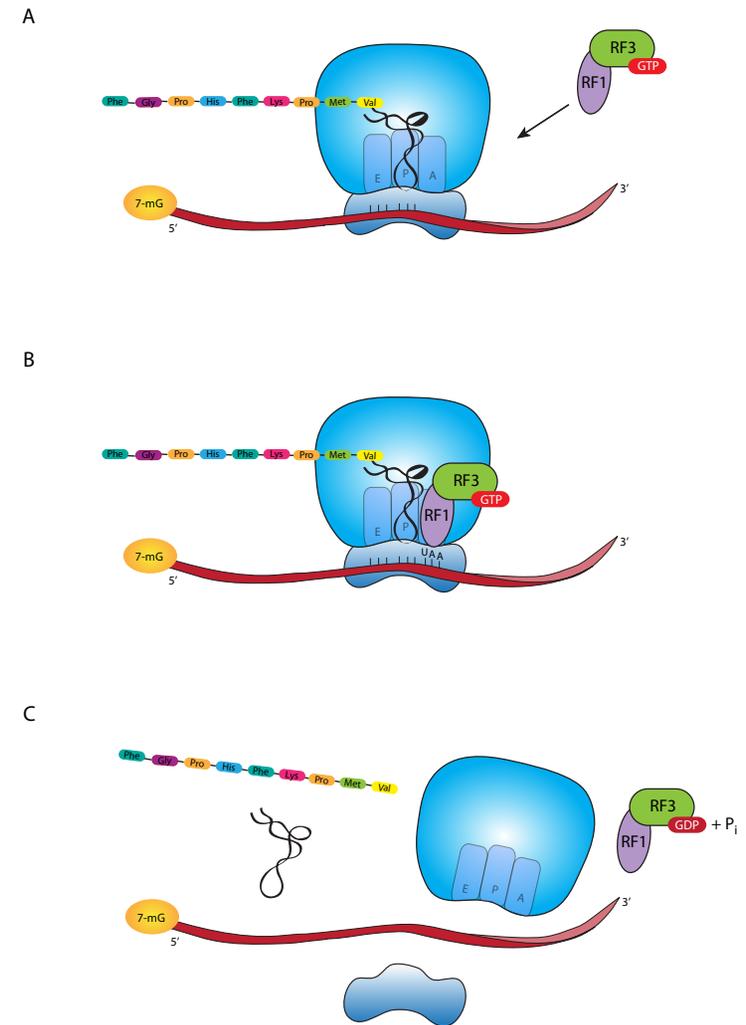


Figure 6. Termination of translation.

For eukaryotes, each mRNA encodes one and only one gene (as opposed to multi-gene transcripts such as operons), so there isn't much question of which AUG is a start codon, and which are just regular methionines. Therefore, there is no requirement for a Shine-Dalgarno sequence in eukaryotes. The small ribosomal subunit, accompanied by eukaryotic initiation factors eIF-3, eIF-2, and Met-tRNA_i, together known as the ternary complex, binds to eIF-1A. Meanwhile, eIF-4A, -4B, -4E, and -4G bind to the 5' (7-methylguanosine) cap of the mRNA (fig. 7A). The small subunit complex and the eIF4/mRNA cap-binding complex interact to form the 43S complex, which then begins scanning the mRNA from 5' to 3' looking for the first AUG.

Once the 43S scanning complex has found the start codon, the initiation factors drop off, and the large ribosomal subunit arrives. The large ribosomal subunit has bound eIF-6, which prevents it from reassociating with small subunits, and its removal is required first. Another factor, eIF-5 enters the scene during the coupling process between the large and small ribosomal subunits, and hydrolysis of an eIF-5-attached GTP is required to complete the docking of the subunits and the formation of a complete functional ribosome on the mRNA.

Elongation is functionally the same as in prokaryotes except that the functions of EF-Tu is taken care of by EF-1 α , also with hydrolysis of GTP. EF-2 is the eukaryotic analog of EF-G, and utilizes GTP hydrolysis for translocation of the ribosome. Termination uses eukaryotic homologues of the release factors, though eRF-1 takes the place of both prokaryotic RF-1 and RF-2.

Although polyribosomes (aka polysomes) can form on both prokaryotic and eukaryotic mRNAs, eukaryotic polysomes have an additional twist. Technically, a polysome is simply an mRNA with multiple ribosomes translating it simultaneously, but in eukaryotes, the polysome also has a unique morphology because it utilizes PABPI, or poly-A binding protein. This protein not only binds to the 3' poly-A tail of an mRNA, it also interacts with the eIF-4 initiation factors, which thus loops the mRNA into a circular shape. That way, once the ribosome reaches the end of the gene and releases from the mRNA, it is physically near the beginning of the mRNA to start translating again.

Usually, but not always, the first AUG is the start codon for eukaryotic genes. However, the context of the AUG matters, and it is a much stronger (i.e. more frequently recognized and used) start codon if there is a purine residue (A or G) at -3 and a G at +4. See Kozak, M., *Biochimie* 76: 815-821, 1994.

Regulation of Translation

Gene expression is primarily regulated at the pre-transcriptional level, but there are a number of mechanisms for regulation of translation as well. One well-studied animal system is the iron-sensitive RNA-binding protein, which regulates the expression of genes involved in regulating intracellular levels of iron ions. Two of these genes, ferritin, which safely sequesters iron ions inside cells, and transferrin, which transports iron from the blood into the cell, both utilize this translational regulation system in a feedback loop to respond to intracellular iron concentration, but they react in opposite ways. The key interaction is between the iron response elements (IRE), which are sequences of mRNA that form short stem-loop structures, and IRE-BP, the protein that recognizes and binds to the IREs. In the case of the ferritin gene, the IRE sequences are situated upstream of the start codon. When there is high iron, the IRE-BP is inactive, and the stem-loop structures are melted and overrun by the ribosome, allowing translation of ferritin, which is an iron-binding protein. As the iron concentration drops, the IRE-BP is activated and binds around the IRE stem-loop structures, stabilizing them and preventing the ribosome from proceeding. This prevents the production of ferritin when there is little iron to bind.

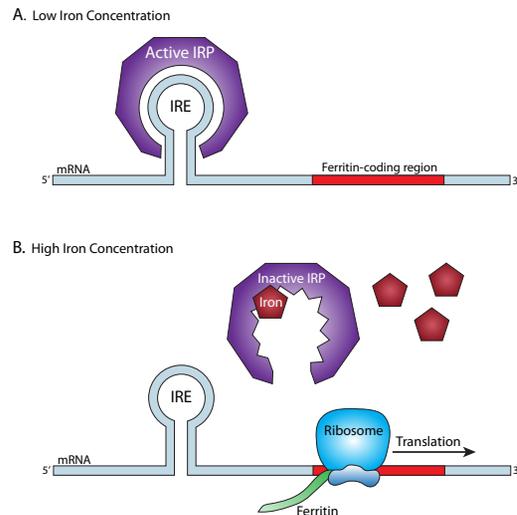


Figure 8. Pre-translational control of gene expression by iron-response protein (IRP), which binds to either the iron-response element (IRE), unless it has bound iron.

Transferrin also uses iron response elements and IRE-binding proteins, but in a very different mechanism. The IRE sequences of the transferrin gene are located downstream of the *stop* codon, and play no direct role in allowing or preventing translation.

However, when there is low intracellular iron and there is a need for more transferrin to bring iron into the cell, the IRE-BP is activated as in the previous case, and it binds to the IREs to stabilize the stem-loop structures. In this case; however, it prevents the 3' poly-A tail degradation that would normally occur over time. Once the poly-A tail is degraded, the rest of the mRNA is destroyed soon thereafter. As mentioned in the transcription chapter, the longer poly-A tails are associated with greater persistence in the cytoplasm, allowing more translation before they are destroyed. The IRE-BP system in this case externally prolongs the lifetime of the mRNA when that gene product is needed in higher amounts.

Since mRNA is a single-stranded nucleic acid and thus able to bind complementary sequence, it is not too surprising to find that one of the ways that a cell can regulate translation is using another piece of RNA. Micro RNAs (miRNAs) were discovered as very short (~20 nucleotides) non-protein-coding genes in the nematode, *C. elegans*. Since their initial discovery (Lee et al, *Cell* 75: 843-54, 1993), hundreds have been found in various eukaryotes, including humans. The expression pattern of the miRNA genes is highly specific to tissue and developmental stage. Many are predicted to form stem-loop structures, and appear to hybridize to 3'-untranslated sequences of mRNA thus blocking initiation of translation on those mRNA molecules. They may also work through a mechanism similar to the siRNA discussed below, but there is clear evidence that mRNA levels are not necessarily altered by miRNA-directed translational control.

Another mechanism for translational control that uses small RNA molecules is RNA interference (RNAi). This was first discovered as an experimentally induced repression of translation when short *double-stranded* RNA molecules, a few hundred nucleotides in length and containing the same sequence as a target mRNA, were introduced into cells. The effect was dramatic: most of the mRNA with the target sequence was quickly destroyed. The current mechanistic model of RNAi repression is that first, the double-stranded molecules are cleaved by an endonuclease called *Dicer*, which cleaves with over-hanging single-stranded 3' ends. This allows the short fragments (siRNA, ~20nt long) to form a complex with several proteins (RISC, RNA-induced silencing complex). The RISC splits the double-stranded fragments into single strands, one of which is an exact complement to the mRNA. Because of the complementarity, this is a stable interaction, and the double-stranded region appears to signal an endonuclease to destroy the mRNA/siRNA hybrid.

The final method of controlling levels of gene expression is control after the fact, i.e., by targeted destruction of the gene product protein. While some proteins keep working until they fall apart, others are only meant for short-term use (e.g. to signal a short phase in the cell cycle) and need to be removed for the cell to function prop-

MicroRNAs are currently under investigation for their roles as either oncogenes or tumor suppressors (reviewed in Garzon et al, *Ann. Rev. Med.* 60: 167-79, 2009). Approximately half of known human miRNAs are located at fragile sites, breakpoints, and other regions associated with cancers (Calin et al, *Proc. Nat. Acad. Sci. (USA)* 101: 2999-3004, 2004). For example, miR-21 is not only up-regulated in a number of tumors, its overexpression blocks apoptosis - a necessary step to allow abnormal cells to continue to live and divide rather than die out. Conversely, miR-15a is significantly depressed in some tumor cells, and overexpression can slow or stop the cell cycle, even inducing apoptosis.

erly. Removal, in this sense, would be a euphemism for chopped up and recycled. The ubiquitin-proteasome system is a tag-and-destroy mechanism in which proteins that have outlived their usefulness are polyubiquitinated. Ubiquitin is a small (76 amino acids, ~5.6 kDa), highly conserved (96% between human and yeast sequences) eukaryotic protein (fig. 9) that can be attached to other proteins through the action of three sequential enzymatic steps, each catalyzed by a different enzyme.

E1 activates the ubiquitin by combining it with ATP to make ubiquitin-adenylate, and then transfers the ubiquitin to itself via a cysteine thioester bond. Through a trans(thio)esterification reaction, the ubiquitin is then transferred to a cysteine in the E2 enzyme, also known as ubiquitin-conjugating enzyme. Finally, E3, or ubiquitin ligase, interacts with both E2-ubiquitin and the protein designated for destruction, transferring the ubiquitin to the target protein. After several rounds, the polyubiquitinated protein is sent to the proteasome for destruction.

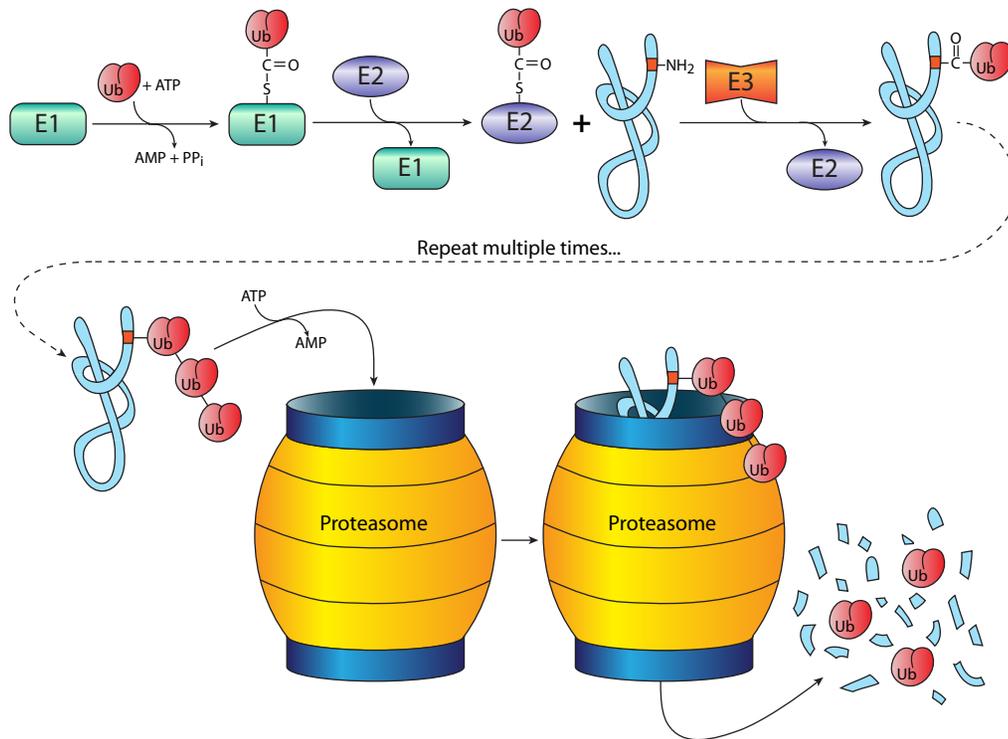
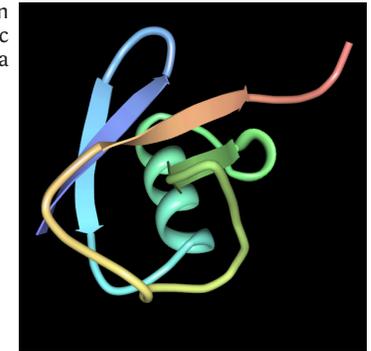


Figure 10. Polyubiquitination of a targeted protein (blue) requires three ubiquitinating enzymes, E1, E2, and E3. Once tagged, the protein is positioned in the proteasome by binding of the polyubiquitin tail to the outer surface of the proteasome. The proteasome then cleaves the protein into small polypeptides.

Figure 9. Ubiquitin. This 3D representation was generated from the file 1ubi (synthetic human ubiquitin) in the RCSB Protein Data Bank



Mutations in E3 genes can cause a variety of human medical disorders such as the neurodevelopmental disorders Angelman syndrome, Hippel-Lindau syndrome, or the general growth disorder known as 3-M syndrome. Mechanisms linking malfunction in ubiquitination pathways and symptoms of these disorders are not currently known.

Proteasomes are very large protein complexes arranged as a four-layered barrel (the 20S subunit) capped by a regulatory subunit (19S) on each end. The two outer rings are each composed of 7 α subunits that function as entry gates to the central rings, each of which is composed of 7 β subunits, and which contain along the interior surface, 6 proteolytic sites. The 19S regulatory units control the opening and closing of the gates into the 20S catalytic barrel. The entire proteasome is sometimes referred to as a 26S particle.

A polyubiquitinated protein is first bound to the 19S regulatory unit in an ATP-dependent reaction (the 19S contains ATPase activity). 19S unit opens the gates of the 20S unit, possibly involving ATP hydrolysis, and guides the protein into the central proteolytic chamber. The protease activity of proteasomes is unique in that it is a threonine protease, and it cuts most proteins into regular 8-9 residue polypeptides, although this can vary.

As we will see in the cell cycle chapter, proteasomes are a crucial component to precise regulation of protein functions.

ER, GOLGI AND VESICLES :

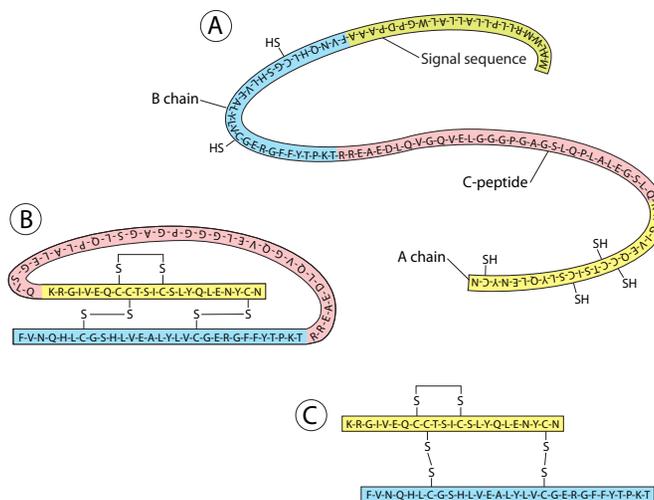
Post-translational Processing and Vesicular Transport

Once a polypeptide has been translated and released from the ribosome, it may be ready for use, but often it must undergo post-translational processing in order to become fully functional. While many of these processes are carried out in both prokaryotes and eukaryotes, the presence of organelles provides the need as well as some of the mechanisms for eukaryote-specific modifications such as glycosylation and targeting.

Proteolytic Cleavage

The most common modification is proteolytic cleavage. Some of the pre-cleavage polypeptides are immediately cleaved, while others are stored as inactive precursors that can be activated very quickly, on a timescale of seconds to minutes, as compared to having to go through transcription and translation, or even just translation. Interestingly, though methionine (Met) is universally the first amino acid of a newly synthesized polypeptide, many proteins have that methionine cleaved off (also true for some prokaryotic f-Met).

Figure 1. Proteolytic processing is necessary to make biologically active insulin. (A) The linear protein contains a signal sequence, which is cleaved after the protein enters the ER, an A chain, a B chain, and a C-peptide. (B) Inside the ER, the proinsulin (insulin precursor) folds and disulfide bonds form between cysteines. (C) Finally, two cleavages release the C peptide, which leaves the A and B chains attached by the disulfide bonds. This is now active insulin.



Using this book: This book is designed to be used in both introductory and advanced cell biology courses. The primary text is generally on the left side of the vertical divider, and printed in black. Details that are usually left to an advanced course are printed in blue and found on the right side of the divider. Finally, additional biomedically relevant information can be found in red print on either side of the divider.

Activation of proteins by cleavage of precursors is a common theme: the precursor protein is termed a proprotein, and the peptide that is cleaved off of it to activate the protein is called the propeptide. Among the better known examples of proteins that are derived from proproteins are the hormone insulin, the cell death protein family of caspases, and the Alzheimer-associated neural protein β -amyloid. Insulin is an interesting example (fig. 1) in mammals: preproinsulin (inactive as a hormone) is first translated from the insulin mRNA. After a cleavage that removes an N-terminal sequence, proinsulin (still inactive) is generated. The proinsulin forms some internal disulfide bonds, and when the final proteolytic action occurs, a substantial chunk (called the C-peptide) is taken out of the middle of the proinsulin. Since the protein was internally disulfide bonded though, the two end pieces remain connected to become the active insulin hormone.

Another interesting protein processing example is that of collagen assembly (fig. 2). As you will read in chapter 13, collagen is a very large secreted protein that provides structure and shock absorbance for the extracellular matrix in animals. You can find it in skin, hooves, cartilage, and various connective tissues. An individual collagen protein is actually a twisted triple-helix of three subunits. The collagen subunits are made as procollagen, and propeptides are lopped off of both N- and C- termini to generate the final protein. However, they are not cleaved off until after the three subunits assemble around one another. In fact, collagen subunits that have already been processed do not assemble into triple-helical proteins. The propeptide sequences are clearly necessary for efficient assembly of the final protein complex.

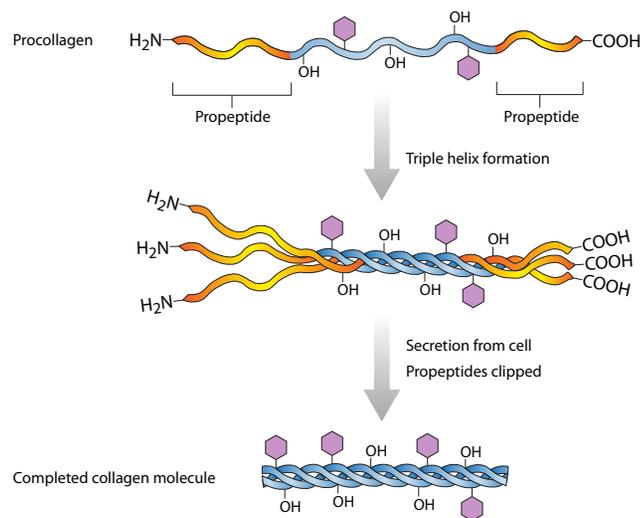


Figure 2. Processing and assembly of procollagen into collagen.

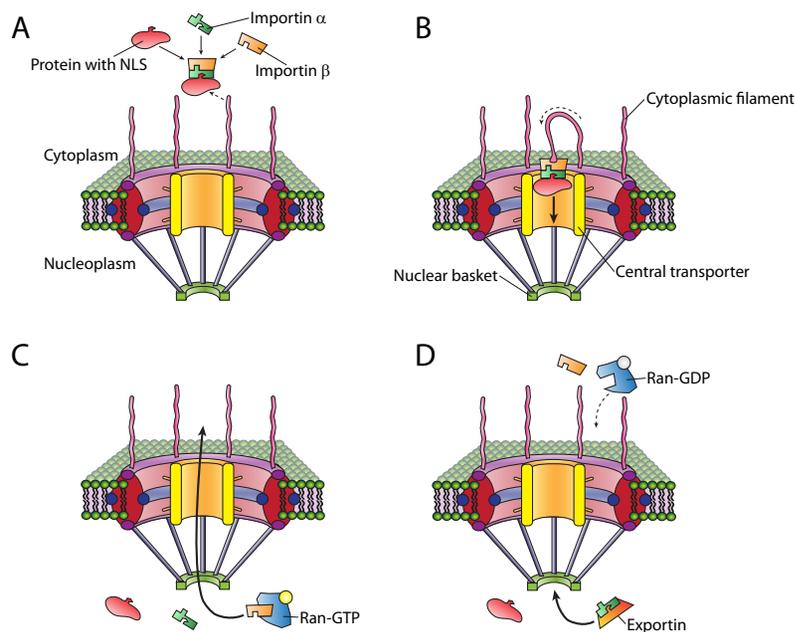
Protein Trafficking

The idea that propeptide sequences have important functions in protein maturation beyond just keeping them from being active is not exclusive to assembly. A major class of cleaved peptide sequences is signal peptides. Signal peptides direct the protein from the cytoplasm into a particular cellular compartment. In the case of prokaryotes, this essentially means the cell membrane, but for eukaryotes, there are specific signal peptides that can direct the protein to the nucleus, to the mitochondria, to the endoplasmic reticulum, and other intracellular organelles. The peptides are specifically recognized by receptors on the membranes of particular compartments, which then help to guide the insertion of the protein into or through the membrane. Almost all protein synthesis in eukaryotes is carried out in the cytoplasm (with the exception of a few proteins in the chloroplasts and mitochondria), so proteins found in any other compartment or embedded in any membrane must have been targeted and transported into that compartment by its signal sequence.

The nucleus is one such compartment, and examples of the proteins found within include DNA and RNA polymerases, transcription factors, and histones. These and other nuclear proteins have an N-terminal signal sequence known as the NLS, or nuclear localization signal. This is a well-studied pathway that involves a set of importin adapter proteins and the nuclear pore complex (fig. 3). Transport into the nucleus is particularly challenging because it has a double membrane (remember that it is contiguous with the endoplasmic reticulum membrane). Although there are other mechanisms for making proteins that are embedded in the nuclear membrane, the primary mechanism for import and export of large molecules into and out of the nucleus itself is the *nuclear pore complex*. The complex is very large and can be made of over 50 different proteins (nucleoporins, sometimes called nups). The nucleoporins are assembled into a large open octagonal pore through the nuclear membranes. As figure 3 indicates, there are antenna-like fibrils on the cytoplasmic face, and these help to guide proteins from their origin in the cytoplasm to the nuclear pore, and on the nuclear side there is a basket structure. Of course, not all proteins are allowed into the nucleus, and the mechanism for distinguishing appropriate targets is straightforward. The protein must bear a nuclear localization signal (NLS). While in the cytoplasm, an importin- α protein binds to the NLS of a nuclear protein, and also binds to an importin- β . The importin- β is recognized and bound by the nuclear pore complex. The details of the transport mechanism are murky, but phenylalanine-glycine repeats in the nucleoporin subunits (FG-nups) are thought to be involved.

Although this is primarily considered a eukaryotic process given that there are so many potential targets, prokaryotes do have membrane proteins (in fact, some 800 different ones in *E. coli* comprising ~20% of total protein), and they are positioned there with the aid of insertase enzymes such as YidC and complexes such as Sec translocase. The Sec translocase uses a signal recognition particle (SRP) much like that in eukaryotes, and will be discussed later in this chapter when the SRP is introduced. YidC, which has eukaryotic homologues (e.g. Oxa1 in mitochondria), is a 61 kDa transmembrane protein that is placed in the membrane through an SRP-Sec translocase mechanism. Once there, YidC interacts with nascent polypeptides (once they reach ~70 amino acids long) that have begun to interact with the lipids of the cell membrane, and pushes the protein into/through the membrane.

Figure 3. Transport through nuclear pore.



Once the nucleoprotein-importin aggregate is moved into the nucleus, Ran-GTP, a small GTPase, causes the aggregate to dissociate (fig. 3c). The imported protein is released in the nucleus. The importins are also released in the nucleus, but they are exported back out again to be reused with another protein targeted for the nucleus.

Export from the nucleus to the cytoplasm also occurs through the nuclear pore. The Ran-GTP is also a part of the export complex (fig. 3d), and in conjunction with an *exportin* protein and whatever is to be exported, is moved out of the nucleus via the nuclear pore. Once in the cytoplasm, the hydrolysis of GTP to GDP by Ran (activated by Ran-GAP, a cytoplasmic protein) provides the energy to dissociate the cargo (e.g. mRNA) from the exporting transport molecules. The Ran-GDP then binds to importins, re-enters the nucleus, and the GDP is exchanged for GTP.

The nuclear pore is the only transport complex that spans dual membrane layers, although there are coordinated pairs of transport complexes in double-membraned organelles such as mitochondria. The transport proteins in the outer mitochondrial membrane link with transport proteins in the inner mitochondrial membrane to move matrix-bound proteins (e.g. those involved in the TCA cycle) in from the cytoplasm. The complexes that move proteins across the outer membrane are made up of Tom (translocator outer membrane) family of proteins. Some of the proteins will stay em-

The mechanisms of small GTPase activation of other processes will be discussed again in more detail in later chapters (cytoskeleton, signaling). The key to understanding the mechanism is to remember that the GTPase hydrolyzes GTP to GDP, but still holds onto the GDP. Although the GTPase will hydrolyze GTP spontaneously, the GTPase-activating protein, GAP (or Ran-GAP in this case) greatly speeds the rate of hydrolysis. In order to cycle the system back to GTP, the GDP is not re-phosphorylated: it is *exchanged* for a new GTP. The exchange is greatly facilitated by the action of an accessory protein, the guanine nucleotide exchange factor (GEF), in this particular case, a Ran-GEF.

bedded in the outer membrane: they are processed by a SAM (sorting and assembly machinery) complex also embedded in the outer membrane). Meanwhile, others continue to the Tim (translocator inner membrane) proteins that move them across the inner membrane. As with the nuclear proteins, there is a consensus signal sequence on mitochondrial proteins that is bound by cytosolic chaperones that bring them to the Tom transporters. As shown in the table below, there are signal sequences/propeptides that target proteins to several other compartments.

Nucleus	–PPKKRKRK–
ER entry	MMSFVSLLLVGIWFATEAEQLTKCEVFQ–
ER retention	–KDEL
Mito. Matrix	MLSLRQSIRFFKPATRTLCSRYLL–
Peroxisome	–SKL
Peroxisome	–RLXXXXXHL

Of particular importance for the rest of this chapter, is the sequence targeting proteins to the endoplasmic reticulum, and by extension, any proteins destined for the ER, the Golgi apparatus, the cell membrane, vesicles and vesicularly-derived compartments, and secretion out of the cell. Here, in addition to an N-terminal signal sequence, the position of secondary internal signal sequences (sometimes called signal patches) helps to determine the disposition of the protein as it enters the ER.

The initial insertion requires recognition of the signal sequence by *SRP*, the signal recognition protein. The SRP is a G-protein and exchanges its bound GDP for a GTP upon binding to a protein's signal sequence. The SRP with its attached protein then docks to a receptor (called the SRP receptor, astoundingly enough) embedded in the ER membrane and extending into the cytoplasm. The SRP usually binds as soon as the signal sequence is available, and when it does so, it arrests translation until it is docked to the ER membrane. Incidentally, this is the origin of the "rough" endoplasmic reticulum: the ribosomes studding the ER are attached to the ER cytoplasmic surface by the nascent polypeptide it is producing and an SRP. The SRP receptor can exist on its own or in association with a *translocon*, which is a bipartite translocation channel. The SRP receptor (SR) is also a GTPase, and is usually carrying a GDP molecule when unassociated. However, upon association with the translocon it exchanges its GDP for a GTP. These GTPs are important because when the SRP binds to the SR, both GTPase activities are activated and the resulting release of energy dissociates both from the translocon and the nascent polypeptide. This relieves the block on translation imposed by the SRP, and the new protein is pushed on through the translocon as it is being synthesized. Once the signal sequence has completely entered the lumen of the ER, it reveals a recognition site for *signal peptidase*, a hydrolytic enzyme that resides in the ER lumen and whose purpose is to snip off the signal peptide.

Prokaryotes also use an SRP homolog. In *E. coli*, the SRP is simple, made up of one protein subunit (Ffh) and a small 4.5S RNA. By comparison, some higher eukaryotes have an SRP comprised of six different proteins subunits and a 7S RNA. Similarly, there is a simple prokaryotic homologue to the SRP receptor, FtsY. An interesting difference is that FtsY generally does not interact with exported proteins, and appears to be necessary only for membrane-embedded proteins. Otherwise, there are many similarities in mechanism for SRP-based insertion of membrane proteins in eukaryotic and prokaryotic species, including GTP dependence, and completion of the mechanism by a translocase (SecYEG in *E. coli*).

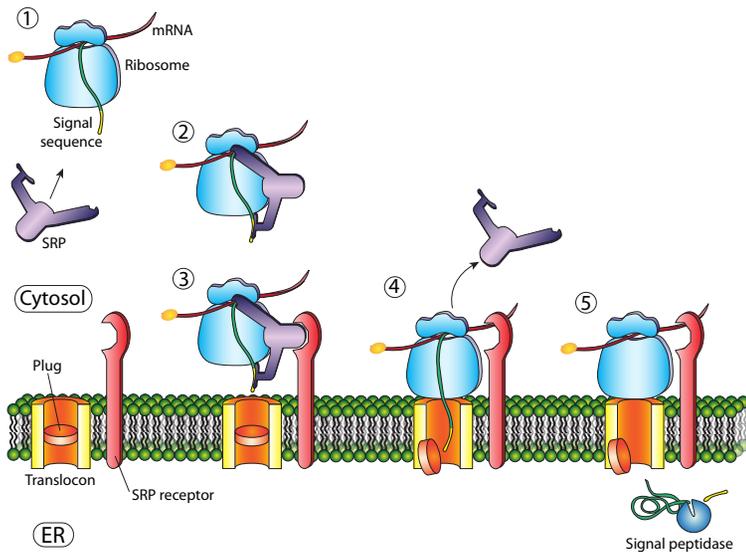
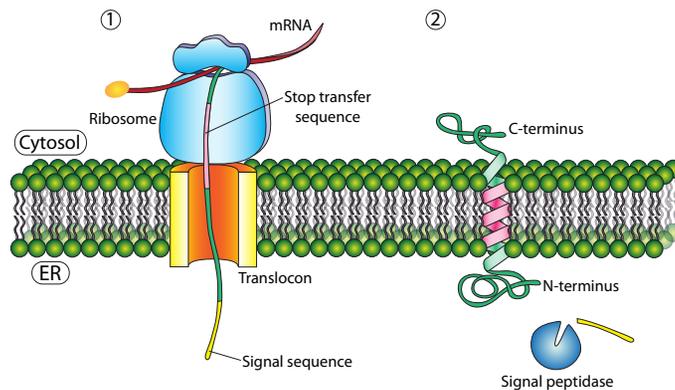


Figure 4. SRP and its receptor SR mediate movement of proteins through the ER membrane. The SRP recognizes the signal sequence and binds to it and the ribosome, temporarily arresting translation. The SRP-polypeptide-ribosome complex is bound by its receptor, SR, which positions the complex on a translocon. Once the ribosome and polypeptide are docked on the translocon, the SRP dissociates, and translation resumes, with the polypeptide moving through the translocon as it is being synthesized.

If that was the only signal sequence in the protein, the remainder of the protein is synthesized and pushed through the translocon and a soluble protein is deposited in the ER lumen, as shown in figure 4. What about proteins that are embedded in a membrane? Transmembrane proteins have internal signal sequences (sometimes called signal patches). Depending on their relative locations, they may be considered either *start-transfer* or *stop-transfer* sequence, where “transfer” refers to translocation of the peptide through the translocon. This is easiest to understand by referring to fig. 5. If

Figure 5. Single-pass transmembrane protein insertion. (1) the signal sequence has allowed the ribosome to dock on a translocon and newly made polypeptide is threaded through until the stop-transfer sequence. (2) The hydrophobic stop transfer sequence gets “stuck” in the membrane, forcing the rest of the polypeptide to stay in the cytoplasm as it is translated.



there is a significant stretch of mostly-uninterrupted hydrophobic residues, it would be considered a stop-transfer signal, as that part of the protein can get stuck in the translocon (and subsequently the ER membrane) forcing the remainder of the protein to remain outside the ER. This would generate a protein that inserts into the membrane once, with its N-terminus in the ER lumen and the C-terminus in the cytoplasm. In a multi-pass transmembrane protein, there could be several start- and stop-transfer hydrophobic signal patches. Building on the single-pass example, if there was another

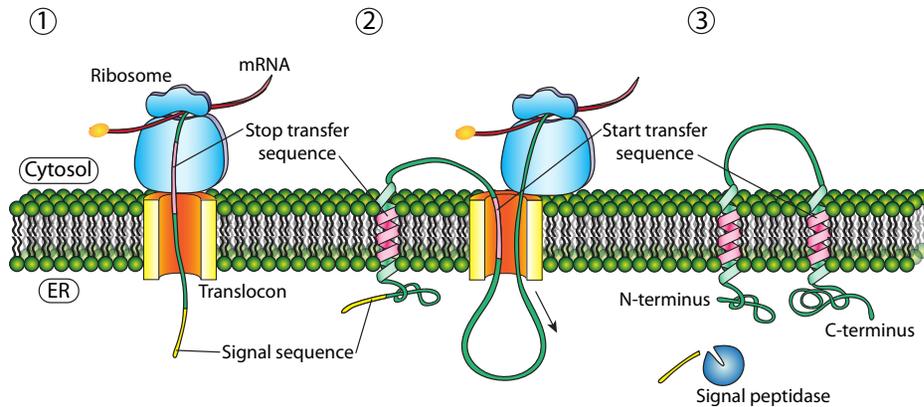


Figure 6. Insertion of 2-pass transmembrane protein.

signal patch after the stop-transfer sequence, it would act as a start-transfer sequence, attaching to a translocon and allowing the remainder of the protein to be moved into the ER. This results in a protein with both N- and C- termini in the ER lumen, passing through the ER membrane twice, and with a cytoplasmic loop sticking out. Of course, the N-terminus could be on the other side. For a cytoplasmic N-terminus, the protein cannot have an N-terminal signal sequence (fig. 7). It has an internal signal patch instead. It plays essentially the same role, but the orientation of the patch means that the N-terminal stays cytoplasmic. The polypeptide translated after the patch is fed into the ER. And just as in the last example, multiple stop- and start-sequences can reinsert the protein in the membrane and change the facing of the next portion.

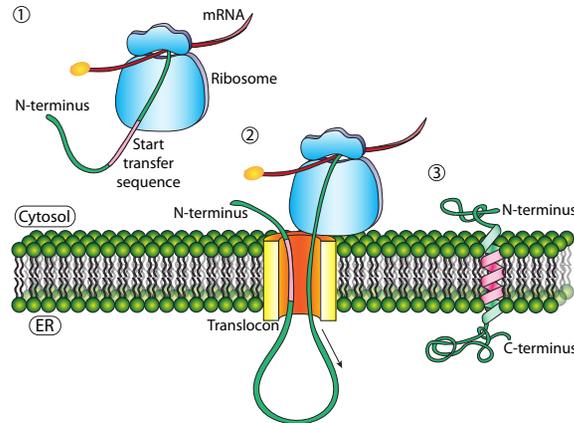


Figure 7. Insertion of a single-pass protein with N-terminus in cytoplasm uses a signal patch but no N-terminal signal.

Protein Folding in the ER

The ER lumen plays four major protein processing roles: folding/refolding of the polypeptide, glycosylation of the protein, assembly of multi-subunit proteins, and packaging of proteins into vesicles. Refolding of proteins is an important process because the initial folding patterns as the polypeptide is still being translated and unfinished may not be the optimal folding pattern once the entire protein is available. This is true not just of H-bonds, but of the more permanent (i.e. covalent) disulfide bonds as well. Looking at the hypothetical example polypeptide, the secondary structure of the N-terminal half may lead to the formation of a stable disulfide bond between the first cysteine and the second cysteine, but in the context of the whole protein, a more stable disulfide bond might be formed between cysteine 1 and cysteine 4. The exchange of disulfide bonding targets is catalyzed by protein disulfide isomerase (PDI). The internal redox environment of the endoplasmic reticulum, is significantly more oxidative than that in the cytoplasm. This is largely determined by glutathione, which is found in a 30:1 GSH:GSSG ratio or higher in the cytoplasm but at nearly 1:1 ratio in the ER lumen. This oxidative environment is also conducive to the disulfide remodeling. It should be noted

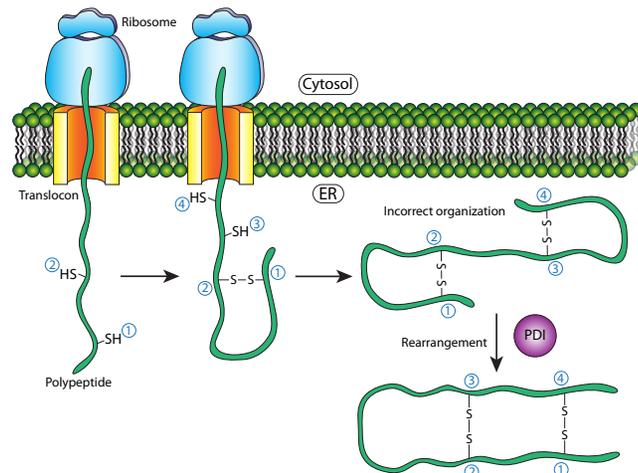


Figure 8. Protein Disulfide Isomerase rearranges disulfide bonds.

The assembly of multisubunit proteins and the refolding of polypeptides are similar in their use of chaperone proteins that help prevent premature folding, sequestering parts of the protein from H-bonding interaction until the full protein is in the ER lumen. This mechanism simply makes finding the thermodynamically optimal conformation easier by preventing the formation of some potential suboptimal conformations. These chaperone proteins bind to the new proteins as they enter the lumen through

that PDI does not choose the “correct” bonding partners. It simply moves the existing disulfide bonds to a more energetically stable arrangement. As the rest of the polypeptide continues to refold, breaking and making H-bonds quickly, new potential disulfide bond partners may move near one another and PDI can again attempt to rearrange the disulfide bonding pattern if the resulting pattern is more thermodynamically stable.

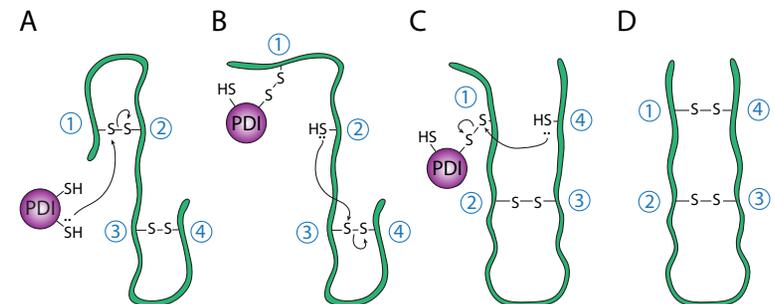
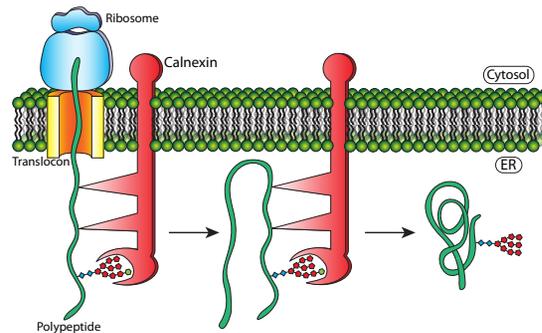


Figure 9. Protein Disulfide Isomerase. This enzyme uses a sulfhydryl group of a cysteine residue as temporary bonding partner in order to break disulfide bonds on the target protein and allow for new ones to form. Note that the formation of a new bond is not directed by PDI, but is instead a stochastic process in which a stronger binding partner displaces the PDI –SH.

Figure 10. Protein folding is optimized in the ER. Proteins such as calnexin can temporarily bind to nascent polypeptides, preventing them from forming secondary structures from incomplete information, releasing the protein for folding once the entire polypeptide has been translated.



the translocon and in addition to simply preventing incorrect bonds that would have to be broken, they also prevent premature interaction of multiple polypeptides with one another. This can be a problem because prior to the proper folding that would normally hide such domains within the protein, the immature polypeptides may have interaction domains exposed, leading to indiscriminate binding, and potentially precipitation of insoluble protein aggregates.

N-linked Protein Glycosylation Begins in the ER

Glycosylation is an important modification to eukaryotic proteins because the added sugar residues are often used as molecular flags or recognition signals to other cells than come in contact with them. There are two types of protein glycosylation, both of which require import of the target polypeptide into the ER. N-linked glycosylation actually begins in the endoplasmic reticulum, but O-linked glycosylation does not occur until the polypeptide has been transported into the Golgi apparatus. Therefore, it is also the case that N-linked glycosylation can (and is) usually beginning as a co-translational mechanism, whereas O-linked glycosylation must be occurring post-translationally. Other major differences in the two types of glycosylation are (1) N-linked glycosylation occurs on asparagine (N) residues within an N-X-S or N-X-T sequence (X is any amino acid other than P or D) while O-linked glycosylation occurs on the side chain hydroxyl oxygen of either serine or threonine residues determined not by surrounding sequence, but by secondary and tertiary structure; (2) N-linked glycosylation begins with a “tree” of 14 specific sugar residues that is then pruned and remodeled, but remains fairly large, while O-linked glycosylation is based on sequential addition of individual sugars, and does not usually extend beyond a few residues.

Technically, N-glycosylation begins before a protein is even being translated, as the dolichol pyrophosphate oligosaccharide (i.e. the sugar “tree” - not an official term, by the way) is synthesized in the ER (fig. 12) without being triggered by translation or

Chaperone proteins can also be found in prokaryotes, archaea, and in the cytoplasm of eukaryotes. These are somewhat similar to each other, and function somewhat differently than the types of folding proteins found in the ER lumen. They are referred to generally as chaperonins, and the best characterized is the GroEL/GroES complex in *E. coli*. As the structure in figure 11 indicates, it is similar in shape to the proteasome, although with a completely different function. GroEL is made up of two stacked rings, each composed of 7 subunits, with a large central cavity and a large area of hydrophobic residues at its opening. GroES is also composed of 7 subunits, and acts as a cap on one end of the GroEL. However, GroES only caps GroEL in the presence of ATP. Upon hydrolysis of the ATP, the chaperonins undergo major concerted conformational changes that impinge on the protein inside, causing refolding, and then the GroES dissociates and the protein is released back into the cytosol.

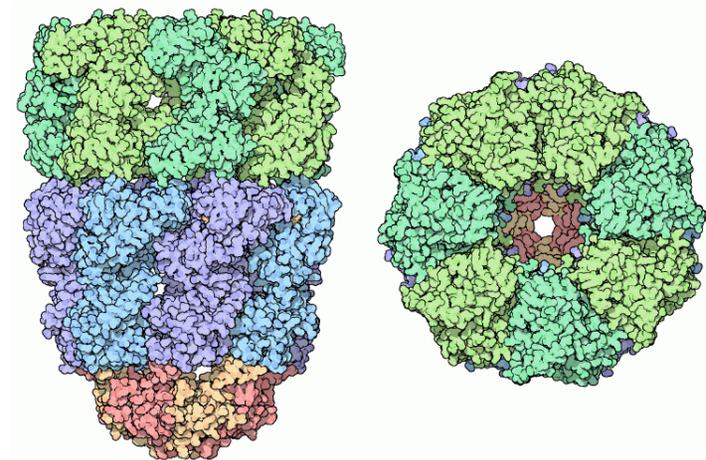


Figure 11. GroEL/GroES complex. The two heptameric rings of GroEL are shown in green and blue/purple. The GroES heptamer (red/yellow) caps the GroEL complex in the presence of ATP. Illustration by D.S. Goodsell, 2002.

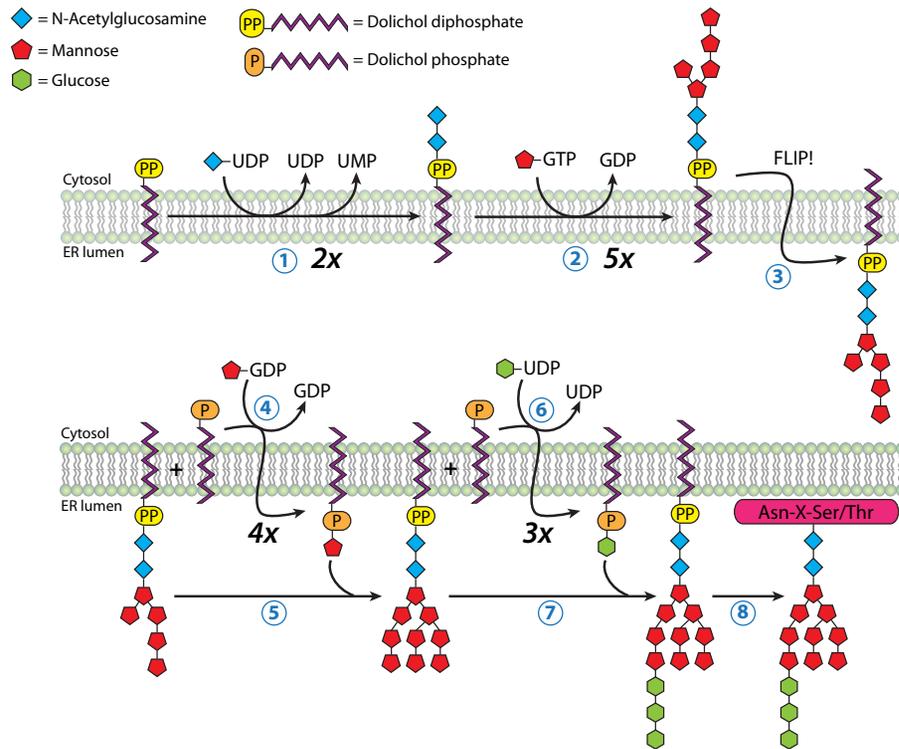


Figure 12. Formation of N-glycosylation “sugar tree” and attachment to protein. Each step is catalyzed by a glycosyltransferase. Note that the sugar substrates are sugar nucleotides, not isolated sugar molecules.

protein entry. Dolichol is a long-chain hydrocarbon [between 14-24 isoprene units of 4+1 carbons] found primarily in the ER membrane, and serves as a temporary anchor for the N-glycosylation oligosaccharide as it is being synthesized and as it waits for an appropriate protein to glycosylate. The oligosaccharide synthesis begins with the addition of two N-acetylglucosamine residues to the pyrophosphate linker, followed by a mannose. From this mannose, the oligosaccharide branches, with one branch receiving three more mannose residues and the other receiving one. So far, all of these additions to the oligosaccharide have been taking place in the cytoplasm. Now the glycolipid is *flipped* inwards to the ER lumen! Once in the lumen, four more mannoses are added, and finally three glucose residues top off the structure.

The enzymes that accomplish the glycosylation are glycosyltransferases specific for both the added sugar residue and the target oligosaccharide. The sugars used by the enzymes are not simply the sugar, but nucleotide sugars - usually a sugar linked to a nucleoside diphosphate, for example, uracil diphosphate glucose (UDP-glucose) or GDP-mannose.

Not all nucleosides are used for this process: sugars have only been found linked to UDP, GDP, and CMP. UDP is the most versatile, binding N-acetylgalactosamine (GalNAc), N-acetylglucosamine (GlcNAc), N-acetylmuramic acid, galactose, glucose, glucuronic acid, and xylose. GDP is used for mannose and fucose, while CMP is only used for sialic acid.

The N-linked oligosaccharide has two physiological roles: it acts as the base for further glycosylation, and it is used as a marker for error-checking of protein folding by the calnexin-calreticulin system (fig. 13). Once the oligosaccharide is attached to the new polypeptide, the process of further glycosylation begins with the action of a glucosidase and that removes two of the glucoses. The last glucose is necessary to help the glycoprotein dock with either calnexin or calreticulin (fig.13, step 1 or 4), which are very similar proteins that have a slow glucosidase activity and associate with a protein disulfide isomerase-like activity. The major difference is that calreticulin is soluble in the ER lumen while calnexin is bound to the ER membrane. Both temporarily hold onto the glycoprotein giving it time to (re)fold and possibly rearrange disulfide bonds, then it removes the glucose, allowing the glycoprotein to continue on its way. Importantly, if the glycoprotein has not been completely folded (step 2a), the enzyme UDP-glucose:glycoprotein glucosyltransferase (GT) recognizes it and adds back the glucose residue (step 3), forcing it to go through the calreticulin/calnexin cycle again in hopes of folding correctly this time. If it has been folded correctly (step 2b), it can be recognized by ER- α -1,2-mannosidase, which removes a mannose, completing the glycosylation modifications in the ER.

The protein disulfide isomerase-like activity comes from ERp57, which is technically a thiol oxidoreductase, but is functionally similar to PDI.

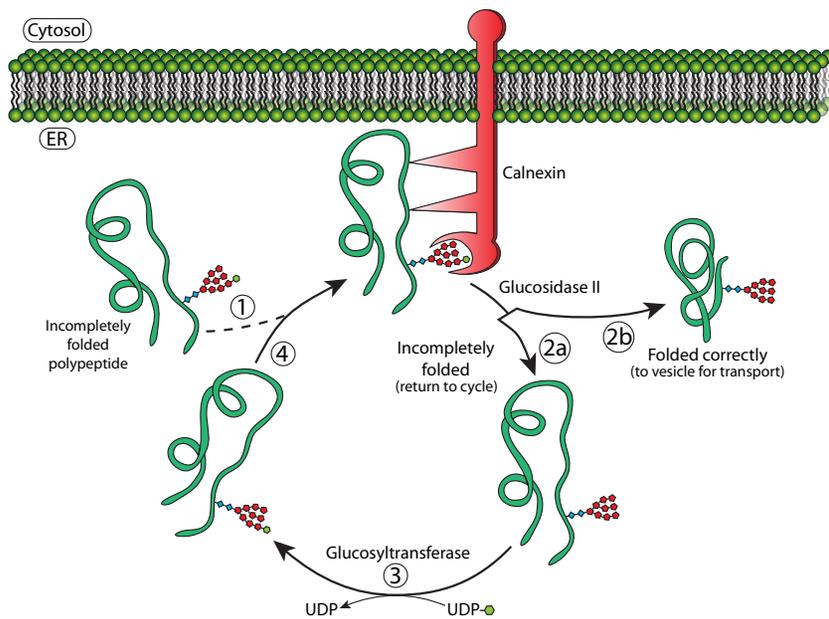


Figure 13. N-glycosylation can be used in error-checking.

Most glycoproteins continue with oligosaccharide remodeling once they have been moved from the ER to the Golgi apparatus by vesicular transport. There, a variety of glycosidases and glycosyltransferases prune and add to the oligosaccharide. Although the glycosylation is consistent and stereotyped for a given protein, it is still unclear exactly how the glycosylation patterns are determined.

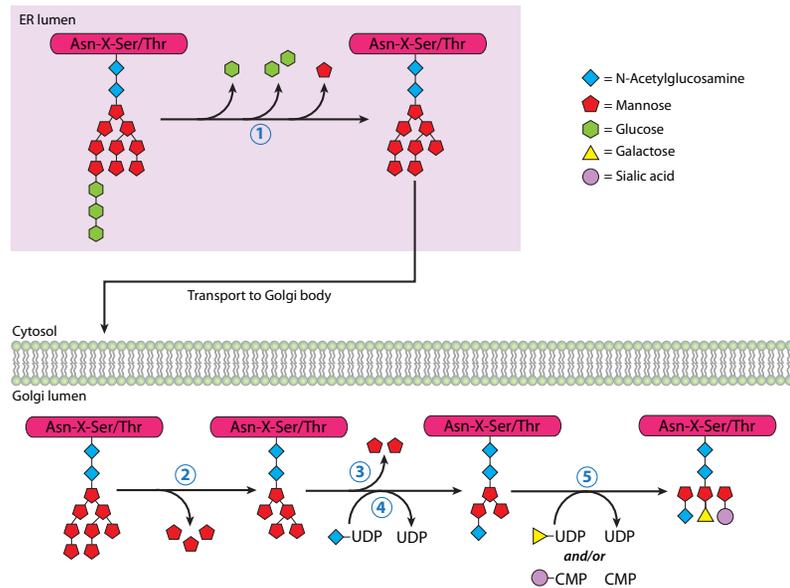


Figure 14. N-linked glycosylation can continue in the Golgi. Sugars may be added and removed in different patterns by glycosyltransferases resident in the Golgi.

O-linked Protein Glycosylation takes place entirely in the Golgi

O-linked glycoproteins begin their glycosylation with the action of the Golgi-specific enzyme, GalNAc transferase, which attaches an N-acetylgalactosamine to the hydroxyl group of a serine or threonine. The determination of which residue to glycosylate appears to be directed by secondary and tertiary structure as previously mentioned, and often occurs in dense clusters of glycosylation. Despite being fairly small additions (usu. <5 residues), the combined oligosaccharide chains attached to an O-linked glycoprotein can contribute over 50% of the mass of a glycoprotein. Two of the better known O-linked glycoproteins are mucin, a component of saliva, and ZP3, a component of the zona pellucida (which protects egg cells). These two examples also illustrate a key property of glycoproteins and glycolipids in general: the sugars are highly hydrophilic and hold water molecules to them, greatly expanding the volume of the protein.

Two common antibiotics, tunicamycin and bacitracin, can target N-linked glycosylation, although their antibiotic properties come from disrupting formation of bacterial cell walls. Tunicamycin is an analogue of UDP-GlcNAc, and inside eukaryotic cells can disrupt the initial oligosaccharide formation by blocking the initial GlcNAc addition to the dolichol-phosphate. Since it can be transported into eukaryotic cells, tunicamycin is not clinically useful due to its toxicity. Bacitracin, on the other hand, is a small cyclic polypeptide that binds to dolichol-PP preventing its dephosphorylation to dolichol-P, which is needed to build the oligosaccharide. Bacitracin is not cell-permeable, so even though it has similar activity to tunicamycin on bacteria by disrupting extracellular glycolipid synthesis needed for cell wall formation, it is harmless to eukaryotes and thus is a useful therapeutic antibiotic.

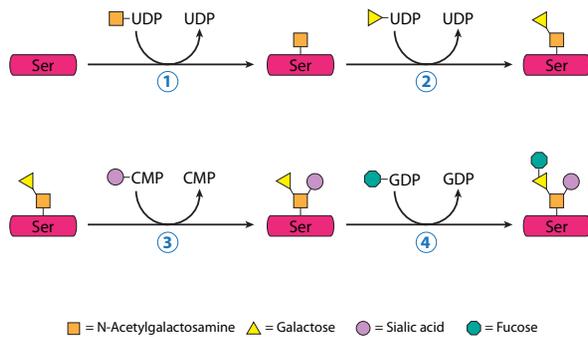


Figure 15. O-linked glycosylation in the Golgi involves attachment of only a few sugars to serine or threonine.

Interestingly, this protective waterlogged shell can mask parts of the protein core. In the case of the cell adhesion molecule, NCAM, which is a highly polysialylated glycoprotein at certain developmental stages and locations, and unglycosylated in others, the naked protein can be recognized as an adhesive substrate while the glycosylated protein can be recognized as a repulsive substrate to other cells. Even in highly glycosylated proteins though, the sugar residues often acts as recognition sites for other cells. For instance, the zona pellucida is very important as a physical barrier that protects the egg, but glycosylated ZP3 also acts as a sperm receptor.

Vesicular Transport

In addition to protein processing, the ER and Golgi also take care of some types of protein transport. Vesicles (membrane-bound bubbles, essentially) pinch off from the ER, Golgi, and other membranous organelles, carrying with them whatever soluble molecules were inside the fluid that was enclosed as well as any molecules embedded in that section of membrane. These vesicles then catch a ride on a molecular motor such as kinesin or myosin, and travel along the cytoskeleton until they dock at the appropriate destination and fuse with the target membrane or organelle. In general, vesicles move from the ER to the cis-Golgi, from the cis- to the medial Golgi, from the medial to the trans- Golgi, and from the trans-Golgi to the plasma membrane or other compartments. Although most movement is in this direction, there are also vesicles that move back from the Golgi to the ER, carrying proteins that were supposed to stay in the ER (e.g. PDI) and were accidentally scooped up within a vesicle.

The formation of vesicles is dependent on coat proteins that will, under proper conditions, self-assemble into spherical cages. When associated with transmembrane proteins, they can pull the attached membrane along into a spherical shape also. The major types of coat proteins used in vesicle formation are COPII, COPI, and clathrin.

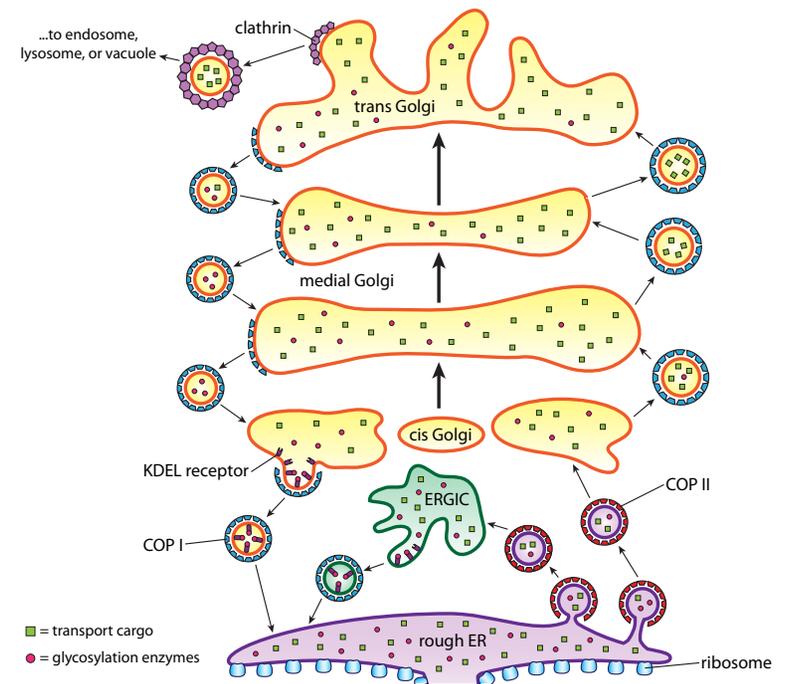


Figure 16. Vesicles bud from the endoplasmic reticulum and merge to form ERGIC, which matures into the cis Golgi, then the medial Golgi, and finally the trans Golgi. Vesicles may also bud from any of these other compartments to other organelles or to the plasma membrane.

COPII coat proteins form the vesicles that move from ER to Golgi. COPI coat proteins are used between parts of the Golgi apparatus as well as to form vesicles going from the Golgi back to the ER. Finally, clathrin is used to form vesicles leaving the Golgi for the plasma membrane as well as for vesicles formed from the plasma membrane for endocytosis.

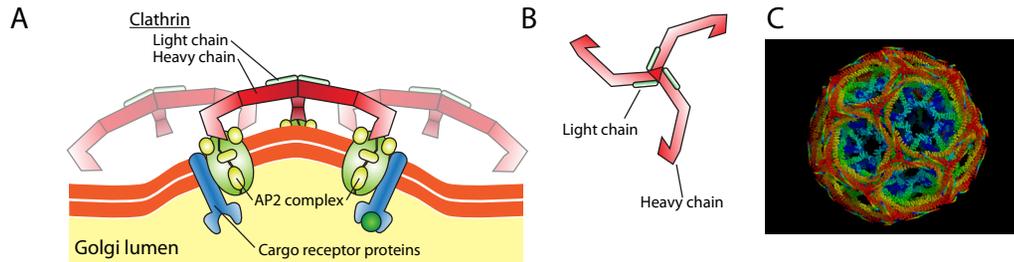


Figure 17. Clathrin. (A) clathrin binds to adapter proteins which are bound to transmembrane cargo receptors, linking the membrane with the clathrin. (B) A single clathrin triskelion is composed of three heavy chains and three light chains. (C) The triskelions self-assemble into a roughly spherical construct without the need for any additional energy or enzymes.

Clathrin (fig. 17) is the best described of the three, and the vesicular coats are made from arrangements of clathrin triskelions (from Gk. meaning three-legged). Each triskelion is composed of three heavy chains joined together at the C-terminus, and three light chains, one associated with each heavy chain. The heavy chains of different triskelions interact along the length of their heavy chain “legs” to create a very sturdy construct. The light chains are unnecessary for vesicle formation, and are thought to help prevent accidental interactions of clathrin molecules in the cytoplasm.

There is significant similarity between the vesicle formation mechanisms using these different coat proteins, beginning with the recruitment of ARF1 (ARF stands for ADP ribosylation factor, which has nothing to do with its function here) to the membrane. This requires the ARNO-facilitated exchange of a GTP for GDP (ARNO is ARF nucleotide binding site opener). Once ARF1 has bound GTP, the conformational change reveals an N-terminal myristoyl group which inserts into the membrane. Both COPI and clathrin-coated vesicles use ARF1 and ARNO, but COPII uses similar proteins called Sar1p and Sec12p.

The ARF1 (or Sar1p) is used to recruit adapter proteins that bind to the “tail” end of membrane-bound receptor proteins. The business end of these receptors binds to cargo molecules that need to be packaged into the vesicle. The adapter proteins act as the link between the membrane (through the receptors) and the coat proteins. For clathrin, the adapter proteins are AP1 for trans-Golgi-derived vesicles and AP2 for endocytic

From the late 1990’s, there has been debate about the origins and stability of the Golgi apparatus. The two primary competing models, both of which are described in many textbooks prior to 2006, were the stable compartment model and the cisternal maturational model. According to the older stable compartment model, the cis, medial, and trans Golgi are each permanent structures with particular characteristics (luminal enzymes, pH, etc). Proteins are shuttled from cis to medial to trans compartments in sequence. The cisternal maturation model, on the other hand, posits that cis Golgi compartments actually mature into medial Golgi, which then mature into trans Golgi - “maturation” being accomplished by import and export of specific proteins characteristic of a particular compartment. Although there was some evidence on either side of the argument, visual demonstration (fluorescent labeling of early- and late- Golgi specific proteins) of cisternal maturation in 2006 (Losev, et al, *Nature* 441: 939-40) confirmed that model.

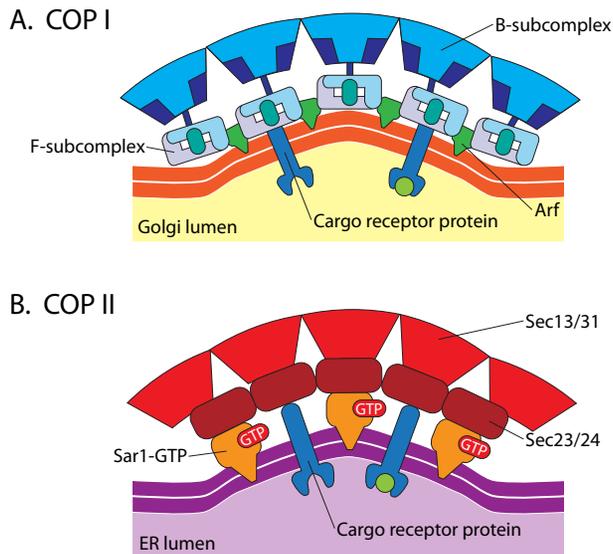


Figure 18. COP coated vesicles.

onal and pentagonal shapes bounded by the clathrin subunits give the vesicle a “soccer ball” look. COP coatamer-coated vesicles are much fuzzier in appearance under EM.

All three types of vesicle coat proteins have the ability to spontaneously associate into a spherical construct, but only the COPI and COPII coated vesicle also spontaneously “pinch off” the membrane to release the vesicle from its originating membrane. Clathrin-coated vesicles require an external mechanism to release the vesicle (fig. 19). Once the vesicle has almost completed, there is still a small stalk or neck of membrane that connects the vesicle to the membrane. Around this stalk, dynamin-GTP molecules aggregate in a ring/spiral construction. Dynamin molecules are globular GTPases that contract upon hydrolysis of GTP. When they associate around the vesicle stalk, each dynamin protein contracts, with the combined effect of constricting the stalk enough that the membrane pinches together, sealing off and releasing the vesicle from the originating membrane.

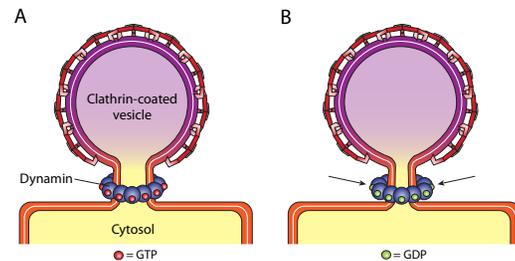


Figure 19. Dynamin monomers, each of which is a GTPase, polymerizes around the neck of the vesicle. When the GTP is hydrolyzed, the dynamin “noose” tightens and pinches off the vesicle.

vesicles. For COPI vesicles, the approximate homologues are the β -, γ -, δ -, and ζ - COPs while the COPII system uses Sec23p and Sec24p.

Finally, the adapters link to the actual coat proteins: clathrin, α - or ϵ - COP, Sec13p and Sec31p. What these proteins all have in common is that spontaneously (i.e. without any requirement for energy expenditure), they self-assemble into cage-like spherical structures. Under the electron microscope, the clathrin-coated vesicles are more sharply defined and the hexagonal and pentagonal shapes bounded by the clathrin subunits give the vesicle a “soccer ball” look.

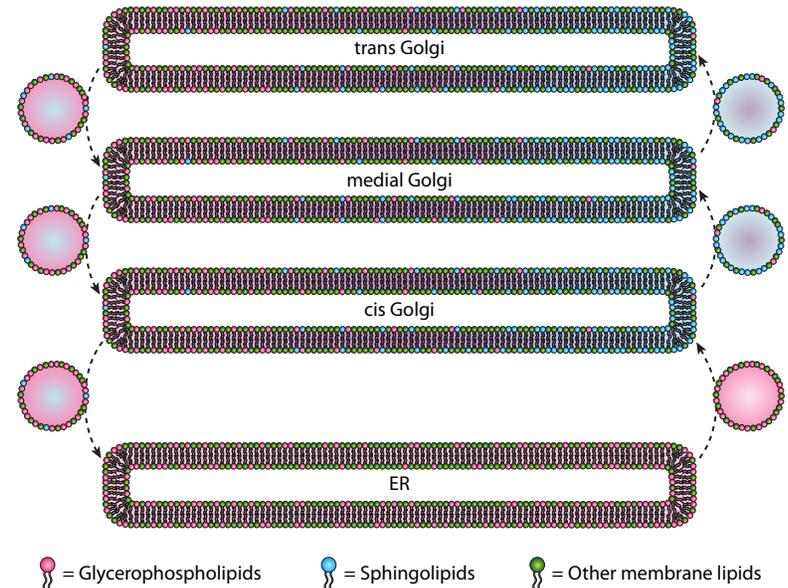


Figure 20. Glycerophospholipids are made primarily in the ER. Although the ER also makes the ceramide precursors for sphingolipids, the sphingolipids are made only in the Golgi.

Although lipids and membranes were discussed in chapter 4, we neglected to discuss the location of their syntheses in eukaryotes. As fig. 20 indicates, the synthesis of certain types of lipids is segregated and exclusive. Glycerophospholipids are primarily formed in the endoplasmic reticulum, although they are also made in mitochondria and peroxisomes. In contrast, sphingolipids are not made in the ER (though their ceramide precursors are) in mammals, the necessary enzymes are found in the lumen of the cis and medial Golgi. There is evidence of anterograde and retrograde vesicular traffic between the various Golgi and ER compartments, which would theoretically indicate a redistribution of lipid types. However, the sphingolipids tend to aggregate into lipid rafts and seem to be more concentrated in anterograde-moving vesicles.

The coat proteins come off shortly after vesicular release. For clathrin, the process involves Hsc70, an ATPase. However, for COPI or COPII coated vesicles, hydrolysis of the GTP on ARF/Sar1p appears to weaken the coat protein affinity for the adapters and initiates uncoating. The GTPase activator is ARF GAP (or Sec23p) and is an integral part of the COP I (or II) coat.

The vesicles carry two categories of cargo: soluble proteins and transmembrane proteins. Of the soluble proteins, some are taken up in the vesicle by virtue of being bound to a receptor. Other proteins just happen to be in the vicinity and are scooped up as the vesicle forms. Occasionally, a protein is taken up that was not supposed to be; for example, PDI may be enclosed in a vesicle forming from the ER. It has little function in the Golgi, and is needed in the ER, so what happens to it? Fortunately, PDI and many other ER proteins have a C-terminal signal sequence, KDEL (Lysine-Aspartic Acid-Glutamic Acid-Leucine), that screams “I belong in the ER.” This sequence is recognized by KDEL receptors inside the Golgi, and binding of the KDEL proteins to the receptors triggers vesicle formation to send them back to the ER.

Secretory vesicles have a special problem with soluble cargo. If the vesicle was to rely simply on enclosing proteins within it during the formation process, it would be difficult to get high concentrations of those proteins. Many secreted proteins are needed by the organism quickly and in significant amounts, so there is a mechanism in the trans Golgi for aggregating secretory proteins. The mechanism uses aggregating proteins such as secretogranin II and chromogranin B that bring together the target proteins in large concentrated granules. These granules work best in the trans Golgi milieu of low pH and high Ca⁺⁺, so when the vesicle releases its contents outside of the cell, the higher pH and lower Ca⁺⁺ breaks apart the aggregates to release the individual proteins.

Finally, there is the question of targeting the vesicles. The vesicles are much less useful if they are tossed on a molecular freight train and dropped off at random. Therefore, there is a docking mechanism that requires a matching of the v-SNARE protein on the vesicle's cytoplasmic surface and the t-SNARE on the cytoplasmic surface of the target membrane. Fusion of the vesicle to the membrane only proceeds if there is a match. Otherwise, the vesicle cannot fuse, and will attach to another molecular motor to head to another, hopefully correct, destination. This process is aided by tethering proteins which initially make contact with an incoming vesicle and draw it close enough to the target to test for SNARE protein interaction. Other proteins on the vesicle and target membranes then interact and if the SNAREs match, can help to “winch” the vesicle into the target membrane, whereupon the membranes fuse. An important rule of thumb to understanding vesicular fusion and also the directionality of membrane proteins and

There is a consistent pH change during the maturation of the Golgi, so that as we go from ER to Golgi, each compartment has a progressively lower (more acidic) luminal pH.

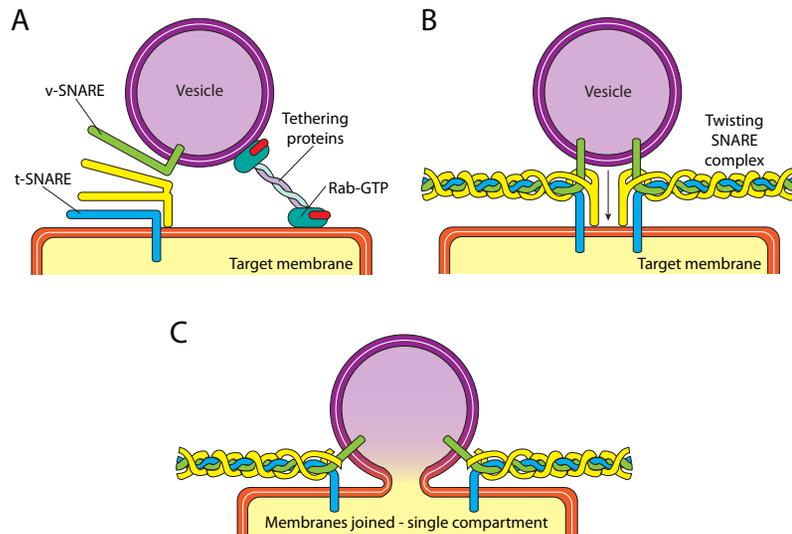


Figure 21. Vesicles first interact with tethering proteins (A), which help bring the vesicle and target membranes close. SNAREs can then interact, and if they match, then they will begin to twist around each other, ratcheting the two membranes closer as they twist.

lipids, is that the cytoplasmic-facing side of a membrane is always going to be facing the cytoplasm. Therefore a protein that is eventually found on the outer surface of the cell membrane will have been inserted into the luminal surface of the ER membrane to begin with.

More specifically, as a vesicle approaches the target membrane, the tethering protein Rab-GTP, which is linked to the target membrane via a double geranylgeranyl lipid tail, loosely associates with the vesicle and holds it in the vicinity of the target membrane to give the SNAREs a chance to work. The v-SNAREs and t-SNAREs now have the opportunity to interact and test for a match. Recently, the SNAREs have been renamed R-SNAREs and Q-SNAREs, respectively, based on conserved arginine and glutamine residues. In addition to these two primary SNAREs, at least one other SNARE is involved, together forming a bundle of four α -helices (four, not three, because at least in the best studied example, one of the SNAREs is bent around so that two of its alpha-helical domains participate in the interaction. The four helices wrap around each other and it is thought that as they do so, they pull the vesicle and the target membrane together.

The tetanus toxin, tetanospasmin, which is released by *Clostridium tetani* bacteria, causes spasms by acting on nerve cells, and preventing neurotransmitter release. The mechanism for this is that it cleaves synaptobrevin, a SNARE protein, so that the synaptic vesicles cannot fuse with the cell membrane. Botulinum toxin, from *Clostridium botulinum*, also acts on SNAREs to prevent vesicle fusion and neurotransmitter release, although it targets different neurons and so has the opposite effect: tetanus is caused by preventing the release of inhibitory neurotransmitters, while botulism is caused by preventing release of excitatory neurotransmitters.

Receptor-mediated Endocytosis

Just as there is vesicular traffic towards the plasma membrane, either for secretion or for incorporation of membrane lipids or proteins, there can also be vesicular traffic from the plasma membrane. Endocytosis is the process by which a coat protein (usually clathrin) on the cytoplasmic side of the plasma membrane, begins to polymerize a coat that draws the membrane with it into a vesicle. However, instead of capturing a bit of ER or Golgi lumen with it, the vesicle contains a little material from outside of the cell. Sometimes endocytosis is initiated internally, perhaps to remove a particular protein from the cell surface (for an example, see trailing edge dynamics in cell motility in the next chapter), but often, the endocytosis is the result of a ligand binding to an extracellular receptor molecule, leading to its activation and subsequent nucleation of a clathrin assembly and vesicle formation.

There are many types of ligands: a nutrient molecule (usually on a carrier protein, as in the examples below) or even an attacking virus which has co-opted the endocytic mechanism to facilitate entry into the cell. The example depicted here is a classic example: endocytosis of cholesterol (via low-density lipoprotein). This illustrates one potential pathway that the receptors and their cargo may take. In the case of cholesterol, the carrier protein is broken down fully, although in the case of transferrin, a serum protein that carries iron in the blood, the carrier protein is just recycled after releasing its transferrin cargo. It is packaged into an exocytic vesicle headed back to the cell surface.

Serum cholesterol is usually esterified and bound by LDL (low density lipoprotein), which then floats about in the bloodstream until it meets up with an LDL receptor on the surface of a cell. When the LDL binds to its receptor, the receptor is activated, and a clathrin-coated vesicle forms around the LDL/receptor complex. LDL receptors tend to aggregate in what are known as clathrin-coated pits – crater-like partial vesicles that already have a small number of polymerized clathrin molecules. The vesicle forms exactly as described previously for Golgi-derived clathrin vesicles: the clathrin self-assembles into a spherical vesicle, and dynamin pinches the vesicle off the cell membrane. This vesicle then fuses with an early endosome, which carries proton pumps in its membrane, causing the environment inside the vesicle to acidify (~pH 6). This acidification can cause conformational shifts in proteins that could, for example, lead to a receptor releasing its ligand, as is the case here with LDL and LDL receptor. The early endosome also functions as a sorting station: the receptor is re-vesicularized and transported back to the plasma membrane. Meanwhile, the LDL is packaged into a different vesicle and heads off for further processing.

The endosomal proton pumps are ATP-driven, Mg^{++} -dependent V-type pump (as opposed to the F-type pump in the mitochondrial inner membrane). Structurally, the two are similar though, and ATP hydrolysis drives the rotary unit, which then powers the movement of protons across the membrane from cytoplasm into endosome.

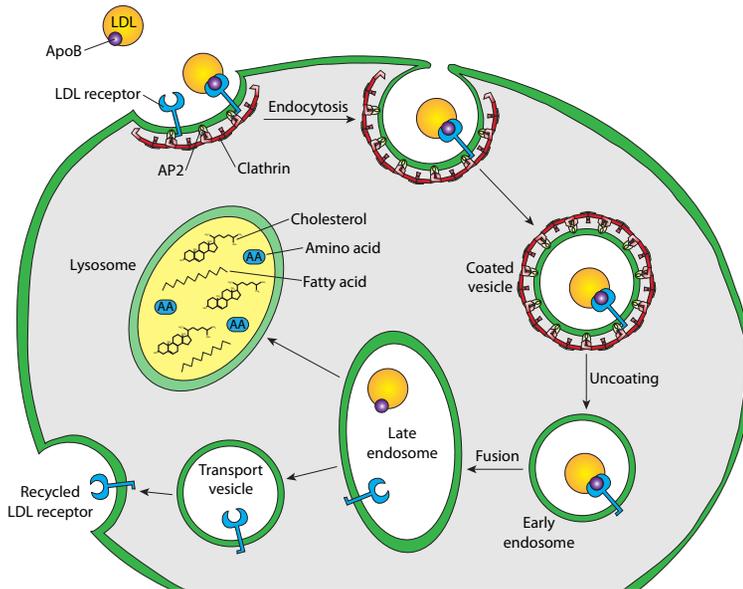


Figure 22. Receptor-mediated endocytosis of cholesterol via low-density lipoprotein.

The endosomal vesicle with the LDL in it next fuses with another acidic, membrane-bound compartment. The lysosome, at pH \sim 5.0, is even more acidic than the endosome, and it also contains a large complement of acid hydrolases - hydrolytic enzymes ranging across substrates (incl. proteases, lipases, glycosidases, nucleases) that operate optimally in acidic conditions, and minimally in the neutral or slightly basic conditions in the cytoplasm. In part, this is a safety mechanism - leakage of digestive enzymes from the lysosome will not result in wholesale digestion of the cell because the enzymes have little or no activity in the cytoplasm. The lysosomal membrane, in addition to having proton pumps to acidify the internal environment, also incorporates many transporter proteins to aid in moving the digestion products of the acid hydrolases out of the lysosome so that the cell can make use of the amino acids, sugars, nucleotides, and lipids that result. Back to our example, that means that the cholesterol esters are broken apart into individual cholesterol molecules, and the lipoprotein is broken down into lipids and amino acids. Interestingly, these transporter proteins are not digested by the lysosomal proteases because they are very heavily glycosylated, which shields potential proteolytic sites from the proteases.

Since it depends greatly on the contents of the endosome(s) that fused with it, the size and contents of lysosomes can vary greatly. In fact, the lysosome may also degrade internal cellular components through the process of *autophagy*. Usually, this is initi-

Lysosomal enzymes are specifically tagged by a mannose-6-phosphate that is added in the cis Golgi. This is a two-step process in which N-acetylglucosamine phosphotransferase adds a phospho-GlcNAc to a mannose residue, connecting via the phosphate group, then a phosphodiesterase removes the GlcNAc, leaving the mannose-6-P. This specifically targets lysosomal enzymes because they all have specific protein recognition sequences that the phosphotransferase binds to before transferring the P-GlcNAc. Although the lysosomal enzymes are tagged in the cis Golgi, they do not sort until the trans Golgi, when mannose-6-P receptors bind to the lysosomal enzymes and form lysosomal vesicles that will bud off and travel to late endosomes and lysosomes to deliver their acid hydrolase payload. Again, the pH change is important: in the somewhat acidic (pH 6.5) environment of the trans Golgi, the receptor binds the mannose-6-P-tagged enzymes, but in the more acidic lysosome, the acid hydrolases are released to do their work.

When one or more acid hydrolases do not function properly or do not make it into the lysosome due to improper sorting, the result is incomplete digestion of the lysosomal contents. This in turn leads to the formation of large inclusions of partially digested material inside the lysosomes. This accumulation of material can be cytotoxic, and genetic disorders that affect the expression or sorting of lysosomal hydrolases are collectively referred to as lysosomal storage diseases. These fall into several categories depending on the types of molecules accumulated.

A common and easily treatable disease of glycosaminoglycan accumulation is Hurler's disease, which can be effectively treated and non-neurological effects even reversed by enzyme replacement therapy. Hurler's others in its class affect a wide variety of tissues because glycosaminoglycans are ubiquitous. On the other hand, because the brain is enriched in gangliosides, lysosomal storage diseases like Gaucher's disease show defects primarily in the CNS. Many lysosomal storage diseases have similar presentation: developmental abnormalities, esp. stunted bone growth, lack of fine facial features, and neuromuscular weakness.

ated under starvation conditions which lead to inhibition of mTor, and subsequent expression of autophagic genes. These then interact with mitochondria and other cellular components, and promote the formation of a double-membraned autophagosome around them. The origin of the membranes is unclear, although the ER is suspected. Finally, the autophagosome fuses with a lysosome, and the acid hydrolases break down the cell parts for energy. A variation on this called microautophagy can also occur, in which the lysosome itself invaginates a bit of cytoplasmic material and internalizes an intralysosomal vesicle that is then broken down.

Finally, it should be noted that the large vacuoles of plant cells are in fact specialized lysosomes. Recall that vacuoles help to maintain the turgor, or outward water pressure on the cell walls that lead to a rigid plant part rather than a limp, wilted one. One of the ways in which this occurs is that the acid hydrolases inside the vacuole alter the osmotic pressure inside the vacuole to regulate the movement of water either in or out.

Another example of receptor-mediated endocytosis is the import of iron into a mammalian cell. As with serum cholesterol, iron is not generally imported into the cell by itself. Instead, it is bound to apotransferrin, a serum protein that binds two Fe^{3+} ions. Once it has bound the iron ions, the apotransferrin is now referred to as transferrin, and it can be recognized and bound by transferrin receptors (TfR) located on the extracellular surface of cell membranes. This initiates receptor-mediated endocytosis just as described above. However, in this case, the lysosome is not involved. As the transferrin and transferrin receptor reach the early endosome, they do not dissociate, but rather the Fe^{2+} releases from the transferrin, and then exits the endosome via DMT1, a divalent metal transport protein to be used in heme groups or other complexes. This leaves the apotransferrin-TfR complex, which is recycled back to the cell membrane via vesicle. Once the vesicle fuses with the extracellular space, the acidity of the endosome is dissipated and the apotransferrin no longer binds to TfR. Apotransferrin can thus go back to its duty of finding iron ions and bringing them back to the cell.

The most severe, I-cell disease (mucopolipidosis type II) occurs when nearly all lysosomal enzymes are missing in the fibroblasts of the affected individual. There is severe developmental delay and early growth failure, neuromuscular problems, and malformations in early skeletal development. The severity of this disorder is due to the almost complete lack of lysosomal enzymes, which is caused by a deficiency of GlcNAc phosphotransferase. Without it, no enzymes are tagged for sorting to the lysosome.

Other relatively common disorders include Tay-Sachs and Niemann-Pick diseases. Tay-Sachs is caused by an accumulation of gangliosides in the brain and is usually fatal by 5 years of age. Niemann-Pick, on the other hand, may manifest as Type A with an even shorter life expectancy, or as Type B, in which symptoms are relatively minor. The major difference is that Type A patients have very little (<5%) of their sphingomyelinase activity, while Type B patients have only slightly less than normal (~90%) activity.