

# Life's Chemical and Biological Processes

AP Biology

1st June 2006

## Contents

<b>1 Cellular Respiration</b>	<b>2</b>
1.1 Introduction . . . . .	2
1.2 Glycolysis . . . . .	3
1.3 Aerobic Respiration . . . . .	6
1.3.1 Pyruvate Oxidation . . . . .	6
1.3.2 Kreb's/Citric Acid Cycle . . . . .	6
1.3.3 Chemiosmosis . . . . .	8
1.4 Anaerobic Respiration . . . . .	9
1.4.1 Alcoholic Fermentation . . . . .	9
1.4.2 Lactic Acid Fermentation . . . . .	11
1.5 ATP Recap and Analysis . . . . .	11
<b>2 DNA and RNA</b>	<b>11</b>
2.1 Introduction . . . . .	11
2.2 DNA Replication . . . . .	15
2.3 Transcription . . . . .	20
2.3.1 Initiation . . . . .	20
2.3.2 Elongation . . . . .	20
2.3.3 Termination . . . . .	21
2.4 RNA Processing . . . . .	21
2.5 Translation . . . . .	22
2.5.1 The Genetic Code and tRNA . . . . .	22
2.5.2 Ribosome (rRNA) structure . . . . .	23

## Preface

Biology is the study and the science of living organisms. The breadth of the subject spans from the macroscopic to the most fundamental level

of chemical behavior and procedural understanding. At the base focus of biology, atomic behavior governs all processes. Passing the atomic scale, molecular interaction takes precedent. Passing the molecular scale, complex chemical processes between molecules make up the recognizable biological processes critical for life to exist.

Detailed in the following pages are the most crucial life processes covered in the AP Biology syllabus. For each process, an overview is given in the form of an introduction, the material is covered “at a glance” and in step by step detail (in many cases, these two sections are meshed together - **IT IS BETTER TO LEARN EVERYTHING!**). Understanding each process “at a glance” is very important for success on the AP exam. On the multiple choice section of the exam, all information given “at a glance” (input, output, purpose) should be well memorized. Understanding each process step by step and in depth is useful, however, not essential for success. That said, when presented with an essay covering one of the explained processes, having step by step knowledge is very useful and will most likely yield full points on that particular question. In addition, knowing each step of each process adds rhyme and reason to the bulk memorization which typically encompasses “at a glance” facts and figures.

When studying the following processes, constantly keep the reoccurring theme in mind: *structure and function*. All of the following processes take place in some structure. To better expedite their processes, these structures have, in almost all cases, adapted significantly. Keep such adaptations, and how they benefit their structure’s processes, in mind at all times!

## 1 Cellular Respiration

### 1.1 Introduction

Cellular respiration occurs in all organisms and is the process that changes food molecules into viable energy for cells. In other words, cellular respiration is the process of changing **glucose to ATP**. Glucose ( $C_6H_{12}O_6$ ) is the classic monosaccharide (1 monomer, 1 simple sugar) carbohydrate (figure 1). ATP (Adenosine triphosphate) is cellular energy currency (figure 2). The energy provided by ATP is based on the high energy bonds between its phosphate groups. When energy is required, the bond between the second and third phosphate group breaks, the ATP loses the third phosphate, and **energy is released**. Once separation is complete, the remaining molecules are ADP

(Adenosine diphosphate) and a lone phosphate (P). Conversely, the opposite of the above reattaches the lone phosphate to ADP. This process (known as **oxidative phosphorylation**, or recreating ATP from ADP and P) **requires energy** and is the end result of cellular respiration.

The process of cellular respiration is based on whether or not oxygen is present. Regardless, cellular respiration begins with a molecule of glucose, which then enters the first step: **glycolysis**. After glycolysis is complete, the resulting compounds are then processed through the **aerobic pathway** (if oxygen is present) or the **anaerobic pathway** (if it is not).

## 1.2 Glycolysis

Glycolysis, itself a 10 step process, is the first step of cellular respiration. **In both eukaryotes and prokaryotes, glycolysis takes place in the cell cytoplasm.** Glycolysis transforms a glucose molecule into 2 pyruvate (or pyruvic acid) molecules. In addition, the following are created indirectly as a result of glycolysis: 4 ATP, 2 NADH, 2 H<sub>2</sub>O. These compounds are formed through the addition of additional compounds to the glycolysis pathway, which promote the pathway further towards completion. Although not direct products of glycolysis, the said compounds are **later** used to synthesize ATP. NOTE: there is a direct ATP gain<sup>1</sup> in glycolysis. While 4 ATP are produced, the net gain after the process is complete is only 2 ATP because 2 ATP are consumed in order to make the process move to completion. Figure 3 references all of glycolysis (depth and detail).

Although memorizing figure 3 is not necessary to achieve success on the AP Biology exam, it will only add to your score. With this in mind, the following tricks make memorizing it in bulk less painful:

1. The first half of glycolysis requires energy input (ATP  $\rightarrow$  ADP); the second half yields energy
2. The enzymes that catalyze steps 1, 3, 7, and 10 are all kinase enzymes
3. *Some* intermediate substance is introduced to the main process at steps 1, 3, 7, and 10
4. Products of the first half of the process can be summarized: glucose, glucose . . . , fructose . . . , fructose . . .

---

<sup>1</sup>Always remember that the goal of cellular respiration is to generate ATP.

Figure 1: Glucose molecule.

Figure 2: ATP molecule.

Figure 3: Glycolysis.

5. Products at steps 6, 7, and 8 are some form of phosphoglycerate

The most important “trick” to employ when memorizing this process is practice<sup>2</sup>. Write it out on napkins, recycled<sup>3</sup> paper, etc. When taking the exam, even if glycolysis is given as an essay question, don’t ever write out the whole process. Reference and know the key information: input, output, intermediates.

### 1.3 Aerobic Respiration

After 2 pyruvate are produced in glycolysis, the pyruvate enter the aerobic respiration pathway if oxygen is present.

#### 1.3.1 Pyruvate Oxidation

Pyruvate oxidation is a 2 step process immediately following glycolysis that transforms pyruvate into Acetyl Coenzyme A (Acetyl CoA). **Pyruvate oxidation takes place in the inner mitochondrial matrix for eukaryotes and in the cell cytoplasm for prokaryotes.** Additional compounds produced in this process are: NADH and CO<sub>2</sub>. Keep in mind that 2 pyruvate are produced during the glycolysis of 1 glucose. As such, the results of pyruvate oxidation after receiving 2 pyruvate are 2 NADH and 2 CO<sub>2</sub>. Figure 4 references all of pyruvate oxidation.

#### 1.3.2 Krebs’s/Citric Acid Cycle

The Krebs’s Cycle (or Citric Acid Cycle, CAC) is an 8 step process that breaks down Acetyl CoA, manufactured during pyruvate oxidation, into an assortment of compounds that are later used to synthesize ATP. **CAC takes place in the inner mitochondrial matrix for eukaryotes and in the cell cytoplasm for prokaryotes.** CAC is a self-perpetuating process. This means that CAC does not make a central product (like pyruvate in glycolysis). Instead, CAC manufactures auxiliary compounds (NADH, etc; similar to those formed peripherally in glycolysis). Specifically, one turn of CAC (when the input is 1 Acetyl CoA) yields 3 NADH, 1 FADH<sub>2</sub>, 1 ATP, and 2 CO<sub>2</sub>. Like in pyruvate oxidation, though, 2 Acetyl CoA are present (manufactured from 2 pyruvate, which in turn is made from 1 glucose). When 2 Acetyl CoA are processed, CAC forms 6 NADH, 2 FADH<sub>2</sub>, 2 ATP, and 4 CO<sub>2</sub>. Figure 5 references all of CAC.

---

<sup>2</sup>Mrs. Flynn Surby will most likely require that this process is memorized for a graded quiz.

<sup>3</sup>Not using recycled paper will wind you up in Hell one day.

Figure 4: Pyruvate oxidation.

Figure 5: Krebs's/Citric Acid Cycle/CAC.

Like glycolysis, memorizing CAC is difficult, not necessary, but very useful. Below are some useful tips for helping to memorize CAC:

1. All of the chemical compounds in each of the 8 steps of CAC are very similar, structurally; as such, memorize the first and then memorize each small change (as opposed to bulk memorizing all 8 structures)
2. In steps 3, 4, 5, 6, 7, and 8, some compound is either being created and/or changed in the system
3. In step 1, Acetyl CoA and Oxaloacetate form citrate, which then forms *isocitrate* (step 2)
4. In step 4, succinyl CoA is formed; in step 5, *succinate* is formed

When working to memorize CAC<sup>4</sup>, don't forget that practice is key!

### 1.3.3 Chemiosmosis

Chemiosmosis (or chemiosmotic theory, oxidative phosphorylation, the electron transport chain) is the final step in manufacturing ATP from glucose in the aerobic respiration pathway. **Chemiosmosis takes place in the inner mitochondrial membrane of eukaryotes and the plasma membrane of prokaryotes.** Chemiosmosis is the process that uses all of the auxiliary compounds (NADH and FADH<sub>2</sub>) made in glycolysis, pyruvate oxidation, and CAC to make ATP. To recap, from all three processes, chemiosmosis has 8 NADH, 2 FADH<sub>2</sub> to process (from 1 glucose). Chemiosmosis is made possible by 5 complexes situated through the inner mitochondrial matrix of the mitochondrion. 4 of the 5 are used to disconnect H<sup>+</sup> ions from the NADH and FADH<sub>2</sub> and shoot those ions into the inner membrane space of the mitochondrion. Specifically, complex 1 deals with NADH (turning it into NAD<sup>+</sup>), complex 2 processes FADH<sub>2</sub> (transforming it into FAD<sup>+</sup>), complex 4 splices H<sub>2</sub> and O<sub>2</sub> (making H<sub>2</sub>O) and complex 5 (**ATP synthase**) phosphorylates ADP into ATP. The first 4 complexes are powered by a lone electron, fired from complex 1, after breaking NADH, which then hits complex 2, forcing it to break FADH<sub>2</sub>, which then fires the electron into complex 3, which then transfers the electron to complex 4 where it is taken in by O and H to help make H<sub>2</sub>O. **Without the electron moving through**

---

<sup>4</sup>Because CAC is a cyclic process, some students have found meditating conducive to memorizing CAC. This is because, since CAC is a circular process, one can surround himself/herself with CAC's different processes (one per sheet of paper), and do the CAC spin.

**the complexes, this process would not function. Without Complex 4 “catching” the fired electron, this process would not function. Without  $O_2$  present, complex 4 would not function.** Therefore, it is complex 4 that makes oxygen essential for the aerobic pathway to complete. Without oxygen, NADH and  $FADH_2$  could not be split from their perspective protons, causing NADH and  $FADH_2$  to accumulate in the mitochondrion, which halts ATP production, causing the cell to die. If oxygen is present, and all  $H^+$  protons are expelled into the inner mitochondrion membrane, a **proton concentration gradient** is formed across the membrane (there is a higher proton concentration on one side of the membrane). This causes, much like water on one side of a dam, the protons to want to flow back to the inner mitochondrial space, which they do through complex 5: ATP synthase. ATP synthase functions as a dam’s turbine, using the flow of  $H^+$  protons through it to phosphorylate ADP with P, making ATP. Figure 6 references all of chemiosmosis.

Chemiosmosis is highly intuitive, and therefore cannot be reduced to tricks when being memorized. Never the less, out of all the processes in cellular respiration, knowing chemiosmotic theory in general is most important on both the multiple choice and free response section of the AP Biology exam.

## 1.4 Anaerobic Respiration

Anaerobic respiration occurs immediately after glycolysis if oxygen is not present (from chemiosmosis, there is no electron acceptor at the end of the electron transport chain). The two types of anaerobic respiration (alcoholic and lactic acid fermentation) are both designed to replenish  $NAD^+$  stores, allowing glycolysis to complete another run, which gives aerobic respiration another chance to activate. Unlike the different pathways of aerobic respiration, which occur one after the other, the anaerobic pathways are a fork. In other words, only one type of fermentation takes place if anaerobic respiration is called into action. Which type is determined based on the organism. NOTE: Knowing the full procedure of anaerobic pathways is not important, and therefore is not discussed.

### 1.4.1 Alcoholic Fermentation

Alcoholic fermentation is the first of two, occurring in plants, fungi, bacteria, and other non-animals. Alcoholic fermentation takes place in two steps, producing  $NAD^+$  and  $CO_2$ . In the first step, 1 pyruvate is converted to 1 acetaldehyde and 1  $CO_2$ . In the second, the acetaldehyde is converted into 1

Figure 6: Chemiosmosis.

ethanol, which, at the same time, forms  $\text{NAD}^+$  from  $\text{NADH}$ . This may seem counter productive because only  $\text{NADH}$  can be used in chemiosmosis to make ATP. Remember, though, that in order to yield the most ATP, pyruvate *must* enter the aerobic pathway. The only way for this to happen is if pyruvate is produced, which requires that glycolysis takes place, which requires  $\text{NAD}^+$ .

### 1.4.2 Lactic Acid Fermentation

Lactic acid fermentation, the second of the two, occurs in animals. This process is single step, converting 1 pyruvate into 1 lactic acid molecule, which simultaneously creates 1  $\text{NAD}^+$  from  $\text{NADH}$ . As in alcoholic fermentation, the  $\text{NAD}^+$  formed promotes glycolysis. The lactic acid produced can, at a later time, be used as a source of energy. Recovering this energy requires oxygen, which, in turn creates an even larger oxygen debt.

## 1.5 ATP Recap and Analysis

The end goal of cellular respiration, aerobic or anaerobic, is to eventually produce ATP. ATP is directly produced in small amounts during glycolysis and CAC. ATP is directly produced in large amounts during chemiosmosis. In chemiosmosis, ATP is generated through the dam-turbine action of ATP synthase that is driven by the proton gradient formed from  $\text{NADH}$  and  $\text{FADH}_2$  being broken apart. Specifically, each processed  $\text{NADH}$  can produce a total of 3 ATP, while each  $\text{FADH}_2$  can produce 2 ATP. The net gain of ATP, **per glucose**, in each process is outlined in table 1.

Yields during anaerobic respiration are maxxed at 2 ATP. This is due to the fact that only ATP directly created by glycolysis can be accessed when oxygen is not present (oxygen must be present for glycolysis' 2  $\text{NADH}$  to count towards an ATP total).

## 2 DNA and RNA

### 2.1 Introduction

DNA (Deoxyribonucleic acid) and RNA (Ribonucleic acid) are responsible for the storage and transport of genetic information. Both DNA and RNA are polymers of the nucleotide monomer. A nucleotide consists of a phosphate, 1 of 2 carbon sugars, and 1 of 5 nitrogen bases. A comparison of the general structure and available base pairs of both DNA and RNA can be found in figure 7.

Figure 7: DNA and RNA.

Process	Source	ATP yield
<b>Glycolysis</b>		
	Direct production	+2
	2 NADH	+6
	Transport of pyruvate into mitochondrion	-2
<b>Pyruvate</b>		
<b>Oxidation</b>	2 NADH	+6
<b>Citric</b>		
<b>Acid</b>	Direct production	+2
<b>Cycle</b>	6 NADH	+18
	2 FADH <sub>2</sub>	+4
<b>Grand total</b>		+36

Table 1: ATP yields: aerobic respiration.

**DNA** serves to store genetic information. DNA is a long polymer of nucleotide monomers (as described above), where each nucleotide consists of a phosphate, a **deoxyribose** sugar, and 1 of 4 nitrogen bases. The 4 nitrogen bases are adenine/guanine (purines) and thymine/cytosine (pyrimidines). A single strand of DNA is assembled by linking phosphate to sugar, phosphate to sugar. A double strand of DNA is created by linking two single strands by their nitrogen base sequence. Nitrogen base linkage is possible through hydrogen bonding and is considered complimentary in nature: **adenine (purine) will only bind with thymine (pyrimidine), guanine (purine) will only bind with cytosine (pyrimidine)** (see figure 8). When a double strand of DNA is formed, it winds in the shape of a double helix of uniform diameter (see figure 7).

The process of recovering the genetic information from DNA and putting it to use in a cell involves several processes and the facilities of **RNA**. RNA, like DNA, is also a polymer of nucleotides. In this case, though, each nucleotide consists of a phosphate, a **ribose** sugar, and 1 of 4 nucleotides. The 4 nitrogen bases are adenine/guanine (purines) and **uracil**/cytosine (pyrimidines). When RNA is manufactured, it follows the same complimentary rules as DNA, however, instead of adenine binding to thymine, adenine binds to uracil (the RNA equivalent of thymine). A single strand of RNA is assembled in the same way as DNA, however, tends to remain single stranded.

Figure 8: DNA nitrogen base pairs: complimentary binding behavior.

Unlike DNA, which is solely designed to store genetic information, RNA takes several forms, all of which function together to transfer the genetic information stored in DNA to the structures that make functional proteins out of the codes described in that DNA. To reiterate: **The purpose of DNA is to code for the thousands of proteins that perform nearly all tasks within a cell.** The specific mechanisms used to complete this task are known as **transcription** and **translation**. As the central dogma of Biology goes:

*DNA is **transcribed** into RNA which is **translated** into protein*

Contributing to this end goal are three types of RNA:

**mRNA (messenger RNA)** serve to transport the genetic information stored in DNA from the cell nucleus to the cell cytoplasm (as DNA only stores genetic information, it is stuck in the cell nucleus). mRNA is referenced in figure 10.

**tRNA (transfer RNA)** flow throughout the cell cytoplasm and attach to select polypeptides (amino acids), carry the polypeptides to mRNA - attached to 2 ribosomal subunits - in the cytoplasm, and add polypeptides in sequence, forming a protein. tRNA is referenced in figure 9.

**rRNA (ribosomal RNA)** make up the subunits that form ribosomes. 2 ribosomal subunits come together in the cell cytoplasm to form a ribosome. rRNA is referenced in figure 10.

All RNA is manufactured in the cell nucleus. Specifically, mRNA is manufactured from a DNA strand in the nucleus during the process of **transcription** and shuttled to the cytoplasm. tRNA and rRNA are constructed in the nucleolus, from additional DNA, and shuttled to the cytoplasm separately. Once in the cytoplasm, mRNA locates and binds to a ribosomal subunit (rRNA), thereby attracting a second ribosomal subunit. Once the two subunits come together, the mRNA is ready to be processed by **translation**.

## 2.2 DNA Replication

DNA replication is the process of making exact copies of DNA. **DNA replication takes place in the cell nucleus.** DNA replication is said to be **semi-conservative** because it involves “unwinding” a given double strand of DNA, using each newly available single strand (template/parent strands) to create new strands (complimentary strands). This idea is referenced in figure 11.

Figure 9: tRNA molecule.

Figure 10: Ribosome (brown: 2 rRNA) with mRNA (purple).

DNA replication is a multiple step process that involves numerous enzymes performing an array of tasks. Such enzymes are listed as follows:

**DNA Polymerase III** moves across the template single strands, adding complimentary base pairs to each parent base pair that it encounters. DNA Polymerase III is responsible for constructing complimentary strands of DNA. **DNA Polymerase only moves in the 5' → 3' direction.**

**Primase** creates RNA primer regions on the lagging DNA strand.

**DNA Ligase** fills the gaps between RNA primer regions and Okozaki fragments on the lagging strand.

**DNA Polymerase I** Replaces RNA, situated at RNA primer regions, with DNA.

**Helicase** unwinds the double strand of template DNA at what is called the replication fork.

**Topoisomerase** keeps the unwound DNA strand from tangling after it has been unwound. Similarly, **single stranded binding proteins** work to ensure that the double strand, once unwound, does not tangle.

**Telomerase** adds telomere sequences to the 3' end of each replicated DNA sequence after DNA replication completes.

DNA replication involves firstly, the double stranded parent DNA and its unwinding. Helicase is responsible for unwinding DNA. Where helicase is present on the DNA strand (where the double stranded DNA is being broken apart) is known as the **replication fork**. The replication fork splits the double stranded DNA into 2 distinct template strands, known as the **leading strand** and the **lagging strand**.

Every strand of DNA has definite direction. If one were to analyze a single strand of DNA, one end would be considered the **3'** end and the other would be the **5'** end. When the leading and lagging strands of DNA are being unwound, the 3' end of the leading strand is always at the replication fork. Likewise, the 5' end of the leading strand is away from the replication fork. The situation for the lagging strand is the opposite: the 3' end is away from the fork and the 5' end is at the fork.

Figure 11: DNA semi-conservative replication.

The direction of the DNA template strands is of the utmost importance during DNA replication. This is because the main acting enzyme, DNA Polymerase III, which constructs the complimentary strands, only moves in the  $5' \rightarrow 3'$  direction. Hence, the leading strand is known as the *leading* strand: DNA Polymerase can move constantly in 1 direction: towards the replication fork, which, for the leading strand, is in the  $3'$  direction.

The same process of creating a complimentary strand for the lagging strand is far different because the DNA Polymerase III must work in the opposite direction as the DNA is being unwound. Imagine this scenario: DNA Polymerase III starts at the replication fork on the lagging strand and starts working away from the fork ( $3'$  direction on the lagging strand). By the time the DNA polymerase reaches the end of the lagging strand, the fork will have moving farther up the DNA strand, and there will be a gap with no complimentary strand attached to the lagging strand (the complimentary strand segment that *is* formed is known as an **Okazaki fragment**). This presents a problem, and is explained step by step as follows:

1. Lagging strand: DNA Polymerase III begins at the replication fork and starts to move away from the fork.
2. As soon as the DNA Polymerase III leaves the fork, the enzyme **Primase** creates a segment of RNA in the place of the complimentary strand (this segment forms the tail of the segment that is soon to be made by the DNA Polymerase III).
3. The DNA Polymerase III moves down the lagging strand, adding nucleotides. When the DNA Polymerase hits the RNA primer tail from the previous Okazaki fragment, the DNA Polymerase III detaches and moves back to the replication fork.
4. (By the time that the DNA Polymerase III returns to the fork, the fork has moved up the double stranded template DNA, leaving space between the last Okazaki fragment's RNA primer and the fork) When the DNA Polymerase III reaches the fork, it begins moving down the lagging strand again, tailed by a new RNA primer region from the resident Primase.

This process rinses and repeats until the replication sequence on both the leading and lagging strands is complete. When the sequence is complete, the leading strand has been replicated completely, and the lagging strand consists of DNA sequences (Okazaki fragments from DNA Polymerase III) that are bound to each other by RNA primer sequences.

When replication completes (i.e. the replication fork reaches the end of the double stranded template DNA), several changes to the lagging strand must take place before the Helicase breaks the now two complete double strands of DNA apart. First, the enzyme **DNA Polymerase I** replaces the RNA primer regions between each Okazaki fragment with corresponding DNA. Second, **DNA Ligase** fills any gaps that exist between each Okazaki fragment and their DNA primer regions (after the RNA has been replaced by DNA). Third, the enzyme **Telomerase** attaches repeat base pair sequences to the 3' end of each DNA strand. These repeat sequences are known as telomeres, and are involved in cell ageing.

## 2.3 Transcription

Transcription is the first half of the central dogma of biology: changing immobile DNA, used to store genetic information, into mobile mRNA, used to carry the said genetic information from the nucleus to the cytoplasm. **Transcription takes place in the cell nucleus.** This process takes place in 3 steps: initiation, elongation, and termination.

Important enzymes involved in transcription are listed as follows:

**RNA Polymerase** creates a complimentary RNA strand from 1 of the 2 template strands that make up a double stranded DNA.

### 2.3.1 Initiation

Initiation marks the beginning of transcription. During initiation, the enzyme **RNA Polymerase** binds to a double stranded DNA at a **promoter region** (a base pair sequence of T - A - T - A nucleotides, called a TATA box) and begins to “unzip” the double stranded DNA. Unzipping DNA entails that the once double stranded structure is being “opened up” so that the RNA Polymerase can process one of the two strands of DNA from the double stranded template.

### 2.3.2 Elongation

Elongation marks the main phase of transcription. During elongation, as the DNA strand has already been unzipped, the RNA Polymerase can begin assembling complimentary RNA nucleotides to the template strand that it is situated on (this process is the RNA equivalent to DNA Polymerase III's work during DNA replication). When moving down the unzipped DNA, RNA Polymerase assembles nucleotides in the 5' → 3' direction, as in DNA

replication, and moves in a continuous fashion. While the DNA strand is partially unzipped, only 1 of its 2 template strands can be transcribed.

RNA Polymerase is different from DNA Polymerase III because it forms only RNA complimentary nucleotides. This entails that RNA Polymerase will form a cytosine nucleotide when presented with a guanine nucleotide (like DNA Polymerase III). Unlike a DNA Polymerase III, however, when presented with an adenine nucleotide, RNA Polymerase will form a **uracil** nucleotide (DNA Polymerase III will form a thymine nucleotide).

### 2.3.3 Termination

Termination is the closing step in transcription. When transcription terminates, the RNA Polymerase stops its progress in creating an RNA complimentary strand, releases the complimentary strand, and closes the unzipped portion of the DNA strand. Termination is triggered when the RNA Polymerase starts processing an A - A - A - A sequence. The complimentary RNA strand is now **mRNA** and will exit the nucleus, once it has been processed (see RNA processing), in search of a ribosome in the cytoplasm (see translation).

## 2.4 RNA Processing

RNA processing is an intermediate step that modifies mRNA between transcription and translation. **RNA processing occurs in the cell nucleus.**

The first half of RNA processing involves connecting a head and a tail to the new mRNA. In this step, the mRNA head (the 5' end of the mRNA) is fastened to a **GTP (guanine triphosphate) cap**. The GTP cap is a guanine purine with additional (3 total) phosphate groups. The GTP cap helps the mRNA attach to the small ribosomal subunit when the mRNA reaches the cytoplasm. In addition to the GTP head cap, a **poly-A tail** is also added to the tail (3' end of mRNA). The poly-A tail is a long sequence of adenine nucleotides (literally A - A - A . . . , the same DNA sequence used to signal the termination phase of transcription). The poly-A tail provides mRNA stability and allows it to pass through the nuclear membrane and into the cytoplasm.

The second half of RNA processing involves the elimination of certain segments of the mRNA strand and reconnecting the segments that are to be

used. Specifically, useless segments are known as **introns**<sup>5</sup> while useful segments are known as **exons**. In other words, this process can be described as “eliminating the introns and splicing the exons.” Eliminating introns is accomplished through the work of **small nuclear ribonucleoproteins**, affectionately known as **snRNPs**<sup>6</sup>. To visualize this process (as internet images are very poor in quality), imagine an mRNA segment with two exons connected on either side to a central intron. To eliminate this intron, 2 snRNPs will connect (1 on either end) to the intron, bend the intron into a lariat shape, disconnect 1 end of the intron, and float the intron away. Because 1 end of the intron is disconnected first, the snRNP is able to fold the exon that the intron was attached to into the other exon (when the 2 exons meet, the second end of the intron is disconnected).

## 2.5 Translation

Translation is the second half of the central dogma of biology: changing the information carried in mRNA, now in the cytoplasm, into a functional polypeptide chain (protein). **Translation takes place in the cell cytoplasm.** This process takes place in 3 steps: initiation, elongation, and termination, and is based on what is called **the genetic code**.

### 2.5.1 The Genetic Code and tRNA

The genetic code, discovered by Nirenburg and Matthaei, describes how an mRNA is read at the ribosome at which it lands in the cytoplasm. Remember, translation reads an mRNA transcript and, depending on the mRNA base pair sequence, creates a protein from it. The question surrounding the genetic code was: “how is the mRNA transcript read? What base pair sequences code for what amino acids (forming a protein)?”

The answer to this question came in triplets, also known as **codons**. Every 3 mRNA base pairs (adjacent to each other) code for a single amino acid, to be added as a link in the chain that is to become a protein. For example, if the 3 base pairs in an mRNA “A - U - G” appear in sequence, the amino acid “methionine” is coded for. Important to know is the fact that AUG is a special case - being the **start codon** - and, in addition to coding for methionine, signals the start of translation (as the TATA box does in transcription). Throughout the genetic code, several other codons perform start codon duties

---

<sup>5</sup>The purpose of the introns is not clear and in this process, they are known as “non-coding segments.”

<sup>6</sup>Pronounce snRNP as it sounds (like smurf with an n and p).

(some only in prokaryotes) and several others perform stop codon (terminates translation, as does the AAA . . . sequence in transcription) duties. The vast majority of codons only code for an amino acid and have no special function.

As a codon is 3 RNA base pairs, there are a possible 64 codons ( $4^3$ ). This presents a problem, as there are only 20 known amino acids. The question: if each codon codes for an amino acid, how are there more possible codons than amino acids? The answer to this inquiry lies in **redundancy**: that more than 1 codon can code for the same amino acid<sup>7</sup>. For example, the codons CUU CUC CUA all code for the amino acid leucine. In all cases, a given codon will *always* code for the same amino acid.

The triplet code structure that makes up mRNA is read by **tRNA**, or transfer RNA. tRNA flows throughout the cell cytoplasm charging itself by picking up a select amino acid. Charging a tRNA molecule requires the amino acid that is due to be attached, ATP as an energy source, and an assistor enzyme. Once charged, the tRNA brings the amino acid to a nearby ribosome complex (with attached mRNA), and deposits the amino acid in the appropriate position relative to the growing amino acid (polypeptide) chain. tRNA structure is referenced in figure 9. Noteworthy information in figure 9 is that tRNA are clover shaped, have an acceptor stem, and have an anticodon loop. The tRNA acceptor stem is the amino acid attachment site and the anticodon loop is where the tRNA's anticodon is compared to currently-being-read codon of mRNA, situated at the ribosomal complex.

### 2.5.2 Ribosome (rRNA) structure

Before the process of translation can be analyzed, it is important to be familiar with the structure of the ribosome, where translation occurs. The ribosome, as described, consists of 2 large rRNA molecules, called **subunits**, and sometimes many smaller structural proteins. Ribosomes can be found free-floating in the cell cytoplasm, bound to the folds of the rough endoplasmic reticulum, or bound to the nuclear envelope (surrounding the nucleus). When translation is not taking place, ribosomes are not one unit, and are found separated by their rRNA subunits (1 subunit is considered the “large” subunit, the other “small”). When mRNA is present, the 2 ribosomal subunits join together and secure the mRNA transcript in preparation for translation.

---

<sup>7</sup>When determining what codon codes for what amino acid, the first 2 letters in the codon triplet are the most important (the third is less important).

Once secured to the mRNA transcript, the 2 ribosomal subunits (specifically the large subunit) form distinct functional sites at which different processes, during translation, take place. The 4 different sites are known as the **T**, **A**, **P**, and **E** sites. In figure 9, the E, P, and A sites are shown (the T site *would* be at the far right of the ribosome). The general function of all sites, in all cases, is listed as follows:

**T (tRNA) site** denotes where incoming *charged* tRNA molecules arrive at the ribosome.

**A (amino acid) site** is the location where the tRNA binds to the mRNA transcript and positions its [the tRNA's] amino acid.

**P (polypeptide) site** is where the growing polypeptide chain is positioned.

**E (exit) site** is the position where uncharged tRNA molecules leave the ribosome to retrieve their next amino acid.