DNA, DNA Replication and Sequencing

Phil McClean
September 2005

The discovery that DNA is the genetic material

DNA is the stuff of life. Embedded in DNA are the sequences that encode for all of the genes that determine what an organism will look like and how that organism will react to its environment. But the discovery that DNA was the master information molecule was spread over nearly 80 years. Early (1866) observations of Ernst Haeckel suggested that nucleus transmitted hereditary information to the next generation. Because of his professional stature in biology, this concept focused the work of others toward the nucleus.

The next milestone in determining the nature of the hereditary information was performed by Friedrich Miescher. He studied pus cells that he collected from bandages he collected from surgeries. The majority of these cells were white blood cells. The cells are distinctive because they are primarily composed of nuclei. In 1871, he called this nuclear material nuclein. Miescher further characterized this material chemically (1874). First, he determined that nuclein was rich in phosphorus. Further analysis revealed that it contained acidic (DNA and RNA) and basic (histone proteins) portions. Even though a link was made between a chemical molecule and the hereditary material, further discoveries were necessary to make a definitive link between DNA and heredity.

The experiments of Fred Griffith (1928) set the stage for the final proof that DNA was the genetic material. Working with lethal and non-lethal strains of the *Streptococcus pneumoniae*, he was able to show that a factor from a lethal strain could convert a non-lethal strain to a lethal strain. Because his factor was able to convert one phenotype to another (the true nature of a gene), Griffith called this factor the transforming principle (Fig. 1). Avery, MacLeod, and McCarty (1944) studied this factor in depth. Their experiments definitively proved that the transforming principle was DNA, and not protein or RNA, the other two constituents in the nucleus (Fig. 2).

In retrospect, the Avery, MacLeod and McCarty experiments were definitive. Others did not think so, and the search continued. The experiments by Hershey and Chase (1952) demonstrated to many that DNA was the hereditary material. Using T2 bacteriophage, they showed that DNA and not protein entered the bacterial cell. Since new T2 phage particles were produced inside the cell, it was concluded that DNA was responsible for its development. Although quite conclusive at the time, scientists now realize that the fact that the DNA was actually contaminated with a bit of protein make these less than compelling experimental results.

Although we focus so much of our attention on DNA, it should be noted that RNA could also be the genetic material. But this is only the case for RNA viruses. Heinz Fraenkel-Conrat (1957) and researchers working with him performed these experiments. Using reconstruction
experiments, they were able to interconvert strains of tobacco mosaic virus by the addition of the RNA isolated from an alternate strain.

**DNA structure**

Now that it was known that DNA was indeed the genetic material, it only remained to determine its structure. That work was completed by James Watson and Francis Crick (1953). They determined that DNA is:

- double-stranded
- the strands are oriented in an anti-parallel manner to each other
- purines nucleotides are opposite pyrimidines nucleotides
- guanine hydrogen bonds with cytosine, while adenine hydrogen bonds with thymine, and hydrogen bonds and hydrophobic bonding between stacked bases stabilize the structure.

Fig. 3 details all of the structural features of the DNA molecule.

It should be noted that Watson and Crick did not perform any experiments. Their conclusions were based on research of others. Erwin Chargaff noted that the concentrations of guanine and cytosine were always equal in DNA. Likewise, the concentrations of adenine and thymine were equal. Rosalind Franklin and Maurice Wilkins used X-ray crystallography to study DNA. From these structures, Watson and Crick deduced that DNA had repeating structures (nucleotides), the DNA was of a constant width and double-stranded. All of this data went into their model of the DNA structure.

**Nucleotide structure**

The DNA that is found within cells is a polynucleotide chain. This chain consists of a string of deoxyribonucleotides. The basic component is the triphosphate deoxynucleotide. Each of these nucleotides consists of the same three components (Fig. 4). These components are:

- deoxyribose sugar
- triosephosphate
- nitrogen base

The nitrogen bases fall into two categories. These are the **purines** and **pyrimidines**. The purines have the same basic two-ring structure. They are distinguished by the side atoms at the carbon 6. The two purines are adenine and guanine. The pyrimidines are single-ring structures. The two pyrimidines, thymine and cytosine, primarily differ by the atom attached to carbon 4. In both cases, the variation in the atoms attached to the ring is important for the interactions that stabilize the double helix.

Hydrogen bonding between the nitrogen bases is a major force in holding the structure together. Watson and Crick correctly predicted that **adenine and thymine form two hydrogen bonds**. The hydrogen donors are nitrogen 3 of thymine (that is accepted by nitrogen 1 of adenine) and the NH₂ group attached to carbon 6 of adenine (that is accepted by the oxygen...
group attached to carbon 4 of thymine). In addition, Watson and Crick observed that **guanine and cytosine form three hydrogen bonds**. The hydrogen donors are the NH$_2$ group attached to carbon 2 of adenine, the hydrogen attached to nitrogen 1 or adenine and the NH$_2$ group attached to carbon 4 of cytosine. These are accepted by the oxygen group attached to carbon 2 of cytosine, nitrogen 3 of cytosine, and the oxygen atom attached to carbon 6 of guanine, respectively.

Watson and Crick also noticed another important feature. The physical size of the two purines was similar, as were the size of the two pyrimidines. Since in each case, they observed a purine hydrogen bonding to a pyrimidine, that meant that the distance between the two side single stranded molecules would be consistent. This conformed to the experimental observations. First, it was now clear why, as observed by Chargaff, the concentrations of guanine and cytosine were the same for each DNA molecule, as were the concentrations of adenine and thymine. They act as partners when base pairing. In addition, this conclusion was consistent with the X-ray defracation studies that showed DNA had a regular width.

**The phosphodiester bond**

Each single strand of DNA is a polynucleotide or a sting of nucleotides. (Fig. 5). The nucleotides are held together by a phosphodiester bond. This bond occurs between the $\alpha$ phosphate of one nucleotide and the 3’ carbon of the partner nucleotide (Fig. 6). The bond is formed by a condensation reaction between the triphosphate nucleotide and terminal nucleotide in the strand. This reaction is performed by the enzyme DNA polymerase. This is a energy consuming reaction. That energy comes from the phosphate group of the incoming nucleotide. When the energy is consumed, a pyrophosphate is released. The result of the phosphodiester bond is the linking of two ribose sugar molecules.

**DNA replication**

DNA replication is the process by which the complete genome of a species is copied. It is an essential process for the maintenance of species integrity. If replication was faulty, then genes would accumulate mutations that could alter the ability of an organism to survive in its environment. If many errors accumulate within the species as a whole, then its success as a species could be compromised.

There are two basic requirements for DNA replication. These are:

- DNA template
- a free 3’-OH group

These requirements are necessary to fulfill the two basic principles of DNA replication. First, DNA replication is **semi-conservative**. This means that the replication product contains one strand from the original DNA molecule and a new strand that used the original stand as template for its synthesis. Therefore that strand serves as the DNA template.
Secondly, replication only proceeds in the 5'-3' direction. For this to occur a free 3’-OH group is necessary. The 3’-OH group is required because it interacts with the incoming deoxynucleotide during the condensation reaction that produces the phosphodiester bond. Once the reaction begins, this group is provided by the last nucleotide added to the growing DNA molecule. The problem, though, arises at the beginning of replication. There is no nucleotide present to be used to make that first phosphodiester bond. That is solved by the addition of a RNA primer. That primer provides the original 3’-OH group in the replication reaction.

The RNA primer is added to the beginning of the replication fork. The fork is created by a series of steps that involves the unwinding of the double-stranded DNA by an enzyme called DNA helicase. Once the template DNA is unwound, it necessary to maintain it in this state while replication is proceeding. This is accomplished by a group of single-stranded DNA-binding proteins. Finally, to prevent DNA from becoming entangled during the unwinding process a family of protein called DNA topoisomerases ensures that the supercoiled state is unwound in a manner that makes the DNA ready for DNA replication.

Once the DNA is unwound, the RNA primer is added. But initially it is only added to one strand. That strand is called the leading strand. It is only added to one strand because the replication fork only moves in the 5'-3' direction. This enables the continuous replication of DNA using the leading strand as a template.

But what about the replication of the other strand? It actually must be replicated in a direction that is opposite of the movement of the replication fork. The discovery of Okazaki fragments lead to our understanding of this dilemma. The other strand of DNA undergoing replication is called the lagging strand. RNA primers are added to this strand, and a short stretch of DNA is replicated. This is repeated continually as the replication forks moves. The result is that lagging strand replication results in a series of short replication fragments. Because of this result, the lagging strand is said to undergo discontinuous replication.

Replication continues until both strands of the replication fork are completely copied. What remains is a clean up process that ensures that the entire DNA molecule consists only of DNA and that the final phosphodiester bonds are completed. This step stitches the whole molecule together. As you are aware, DNA only consists of DNA. What becomes of the RNA primer. These are removed by the 5'-3' enzymatic action of one of the enzymes involved in replication. Once these primers are removed, synthesis continues and these newly generated gaps are filled in. Then the final covalent bond is made to seal the molecule together.

This entire process is demonstrated in Figs. 7a and 7b.

**DNA Polymerases**

DNA replication involves a number of reactions. We have discussed three of the four main ones. The 5'-3’ synthesis reaction adds the bulk of the nucleotides to the growing chain. The second reaction is the removal of the RNA primers in a 5’-3’ direction. These two steps are completed by the enzymes DNA polymerase. The other reaction, the closing of the final covalent bond, is completed by an enzyme called DNA ligase.
As you can imagine, DNA polymerases have been studied in great detail especially the *E. coli* enzyme. The replication steps detailed above for the replication of this bacterium involve two different polymerases: DNA polymerase I (PolI) and DNA polymerase III (PolIII). The bulk of the synthesis, the adding of nucleotides in the 5'-3' direction, is performed by DNA PolIII, a three-subunit protein. This is the 5'-3' synthesis function. DNA PolI, a single peptide protein, also has a DNA synthesis function, but it is not involved in the bulk of DNA synthesis. Its main function is to remove the RNA primers. This is accomplished by the 5'-3' exonuclease function. Its synthesis function is used primarily to fill in the gaps left by the removal of the primers.

Errors occur during DNA synthesis. The most common error is the incorporation of an incorrect nucleotide. This cannot be tolerated if the organism is to maintain sequence integrity from generation to generation. This problem is solved by the final function of DNA polymerases (such as PolI and PolIII): 3'-5' exonuclease activity. Once the enzyme detects an error, it removes several nucleotides (including the error) and then returns to its synthesis function. This process is called proofreading. For DNA PolIII, this function resides on a different subunit (ε) than the one performing the 5'-3' synthesis (α). Both the proofreading and synthesis are provided by the single DNA PolI protein.

Eukaryotes have evolved a larger array of DNA polymerases. Each has a specialized function. The semi-conservative replication of the nuclear DNA is completed by polymerase enzymes α and δ. Mitochondrial DNA is replicated by polymerase γ. Two other polymerases, β and ε, are involved in the repair process. Other than the suite of enzymes involved, replication in eukaryotes involves the same basic steps.

**DNA sequencing**

It is important to understand DNA replication because its principles underlie DNA sequencing. With that basic knowledge now in place, we can now discuss DNA sequencing technology. Why is it important to understand DNA sequencing? The basic unit of information in genomics is the DNA sequence of the genome. Therefore, understanding how sequence information is collected is necessary to fully appreciate genomic data.

In 1977, two different DNA sequencing procedures were described. The chemical sequencing procedure involves the chemical cleavage of DNA molecules in a specific manner that allows the sequences to be read. Because the chemicals were somewhat caustic, the procedure never became widely adapted. Another reason is that the alternate procedure, chain termination (or Sanger) sequencing, was an enzymatic procedure that was relatively safe and easier to perform than chemical procedure. Nearly all sequencing today uses the chain-termination procedure.

The basic concept of the chain-termination procedure is the halting of DNA synthesis by the addition of a specialized deoxyribonucleotide called 2',3' dideoxyribonucleotide. As you remember, all nucleotides have a H atom attached to the 2' carbon and a OH group attached to the 3' carbon. The 3'-OH group is essential for DNA replication. The dideoxynucleotide has
this 3’-OH group replaced with a H atom. Once this nucleotide is inserted into the growing DNA chain, DNA synthesis is terminated. (Thus, the term chain-termination is used.)

When performing a sequencing reaction, these specialized nucleotides are added in low concentrations. These low concentrations ensure that a family of DNA fragments is produced, and each fragment differs in length by one nucleotide (Fig. 8). It is now important to visualize these fragments so that the DNA sequence can be determined. There are two commonly used procedures.

The gel-based system uses four different reactions. In each reaction, only one of the four dideoxynucleotides is included. These nucleotides are labeled with a radioisotope. Therefore each fragment has a radiolabel that is detected after the fragments are separated in a polyacrylamide gel (Fig. 9). Once the fragment image is obtained, it is a simple matter of directly reading the sequencing.

Advances in the development fluorescent labels, the detection of fluorescent labels, and computer-based sequence analysis has greatly accelerated the collection of sequence data. This has lead to a modification of the traditional chain-termination sequencing procedure. The primary difference is that all four dideoxynucleotides are added at low concentrations to a single reaction. The nucleotides in this reaction are unique in that each contains a fluorescent molecule attached to it. And each of these molecules can be detected at a specific wavelength.

The sequencing fragments are loaded onto a capillary electrophoresis system. This system separates the molecule by size; the smaller fragments exit first, the larger fragments migrate through slower. As the molecules exit the system, the fluorescent dye is detected by a laser sensor. That data is then sent to a computer where it is stored and later analyzed. The advantage of this system is that all of the sequence fragments are contained within a single capillary. The capillary electrophoresis systems vary in that they have eight, 16, 96, even 384 capillaries that can simultaneously collect sequence data. Typically, the sequence of a reaction can be collected within two hours. As you can imagine, this system greatly increased the throughput of any sequencing project. This process is depicted for a single capillary in Fig. 10.
Griffith and the Transforming Principle

A. The Concept

The experiments of Griffith and Avery, MacLeod and McCarty are closely related. Griffith developed the concept of the **transforming principle**. The principle was able to transform a non-pathogenic bacteria into a pathogenic strain. Changing phenotype is one of the characteristics of the hereditary material. Griffith called the factor that changed the phenotype the transforming principle. Avery, McCarty, and MacLeod performed a series of experiments that demonstrated **the hereditary materials was DNA**.

Fred Griffith’s experiments provided the experimental platform for Avery, McCarty, and MacLeod to prove the DNA was the genetic material. He worked with the pathogenic bacteria *Streptococcus pneumoniae* that is lethal to mice. But not all types of the bacteria all lethal: type R is non-lethal, whereas type S is lethal. In addition, there are type II an III strains of the bacteria. Each of these can be either R or S. So a Type IIIS strain is lethal, whereas a type IIR is non-lethal.

Griffith was able to show that if you heat kill a Type IIIS strain and injected it into the mouse, the mouse lived. But if you mixed the heat-killed type IIIS material with live type IIR bacteria, the mouse would die. Furthermore, the autopsy showed that the mouse became infected with the Type IIIS strain. These meant that some material from the Type IIIS strain was taken up by the Type IIR strain to convert it into the Type IIIS strain. Griffith termed the material the **transforming principle**.

One feature of the genetic material is its ability to control phenotype. In Griffith’s experiment, the bacterial strains have several phenotypes. The R types are not only non-lethal, and they have a rough (R) appearance on a blood agar plate. The S type are distinct from the R type: they are lethal and have a smooth morphology on the plates. The S types have a polysaccharide capsule that is lacking in the R types. Each capsule type is distinguished using antibodies; the type II capsule is antigenically distinct from the type III. The transformation from type II to type III and the conversion of type R to S are each distinct phenotypic changes. Therefore if the chemical nature of the transforming principle could be determined, then we would know the nature of the genetic material. Avery, MacLeod and McCarty found the answer.

**Figure 1.** The experiment of Griffith that demonstrated the concept of the transforming principle.
Avery, MacLeod and McCarty: DNA Is The Genetic Material

A. The Concept

Avery, MacLeod and McCarty extended the work of Griffith. They worked with his system, but rather than working with mice they only studied the bacterial phenotypes relative to the material from the dead type IIIS. They performed careful analysis and proved that DNA and not protein or RNA was the genetic material.

Rather than work with mice, Avery, MacLeod and McCarty used the phenotype of the Streptococcus pneumoniae cells expressed on blood agar. To ensure, a few potentially live cells did not escape the heat treatment, they also precipitated those cells out of culture using an antibody to the type IIR cells. Finally, they included an enzyme treatment of the material from the heat-killed cells. Each of these enzyme destroyed either proteins (protease), RNA (RNase), or DNA (DNase). These are the three main components of the heat-killed cells. As you can see above, the only treatment that prevented the conversion of the type IIR cells to type IIIS was DNase. This demonstrated conclusively that DNA was the transforming principle and the heredity chemical of life.

Figure 2. The experiment of Avery, MacLeod and McCarty that demonstrated that DNA was the genetic material.
DNA Structure

A. The Concept

DNA has a regular structure. It’s orientation, width, width between nucleotides, length and number of nucleotides per helical turn is constant. All of these features were described by Watson and Crick. Adenine is always opposite thymine, and cytosine is always opposite guanine. The two strands are held together by hydrogen bonds: two bonds between adenine and thymine and three bonds between guanine and cytosine.

One helical turn =
3.4 nm =
10 nucleotides

0.34 nm between nucleotides

This figure describes the general features of B DNA, the most common structure found within a cell. Other forms of DNA also exist. All forms have unique features. These are:

<table>
<thead>
<tr>
<th>Form</th>
<th>Helix Direction</th>
<th>Nucleotides per turn</th>
<th>Helix Diameter</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Right</td>
<td>11</td>
<td>2.3 nm</td>
</tr>
<tr>
<td>B</td>
<td>Right</td>
<td>10</td>
<td>2.0 nm</td>
</tr>
<tr>
<td>Z</td>
<td>Left</td>
<td>12</td>
<td>1.8 nm</td>
</tr>
</tbody>
</table>

Figure 3. The structure of common DNA molecules.
Deoxyribonucleotide Structure

A. The Concept

DNA is a string of deoxyribonucleotides. These consist of three different components. These are the deoxyribose sugar, a phosphate group, and a nitrogen base. Variation in the nitrogen base composition distinguishes each of the four deoxyribonucleotides.

The basic building block is the deoxyribose sugar. This sugar is distinguished because it contains a hydrogen (H) atom at the number 2' carbon. Normal ribose has a hydroxyl (-OH) group at this position.

Attached to the 5' carbon is a triphosphate group. This group is important because in a DNA chain it undergoes a reaction with the 3' OH group to produce polydeoxynucleotide.

The final feature of the molecule is a nitrogen base. These are attached to the 1' carbon. Four bases are possible. Two pyrimidines (thymine and cytosine) and two purines (adenine and guanine). The double stranded DNA molecule is held together by hydrogen bonds. Pairing involves specific atoms in each base. Adenine pairs with the thymine, and guanine pairs with cytosine. These pairings and the atoms involved are shown to the right.

You have probably heard of ATP, the energy molecule. It is the deoxyribonucleotide to which adenine is attached. This molecule serves two very important functions in biological organisms.

Figure 4. The structure of deoxyribonucleotides and base pairing among N bases.
A. The Concept

Each strand of the double-stranded DNA molecule has the same basic structure. It is a series of series of deoxyribonucleotides linked together by phosphodiester bonds.

DNA is a polynucleotide. It consists of a series of deoxyribonucleotides that are joined by phosphodiester bonds. This bond joins the a phosphate group to the 3' carbon of the deoxyribose sugar.

Each strand is complementary to the opposite strand. If one strand has an adenine at a position, its anti-parallel strand would have a thymine at the corresponding position. Likewise, guanine and cytosine would be complementary.

Fig. 5. The single strand structure of DNA.
Making a Phosphodiester Bond/ 
Growing the DNA Chain

A. The Concept

The addition of a new nucleotide to a DNA molecule creates a phosphodiester bond. This requires the DNA chain that is being elongated and a deoxyribonucleotide.

Phosphodiester bonds are formed when a new dideoxynucleotide is added to a growing DNA molecule. During the reaction, a condensation reaction occurs between the $\alpha$ phosphate of the nucleotide and the hydroxyl group attached to the 3' carbon. This reaction is performed by the enzyme DNA polymerase. This is also an energy requiring reaction. The energy is provided by the breaking of the high-energy phosphate bond in the nucleotide. This results in the release of a pyrophosphate molecule.

Figure 6. The formation of the phosphodiester bond that grows the DNA chain.
Steps of DNA Replication (Part 1)

A. The Concept

DNA replication is an essential biological process. Its primary function is to produce new DNA for cell division. The process has several distinct steps that are important to understand. The factors that are absolute requirements for DNA replication to begin are a free 3'-OH group and a DNA template. A RNA primer provides the free 3'-OH group. The DNA to be replicated serves as the template. It is important to remember that all DNA replication proceeds in the 5'-3' direction.

1. The replication fork is formed; RNA primer added.

2. DNA is replicated by the 5'-3' synthesis function of DNA polymerase using the leading strand in a continuous manner.

3. An error occurs during DNA replication.

4. The DNA replication error is removed by 3'-5' exonuclease function of DNA polymerase.

Notes on *E. coli* replication:

**DNA Polymerase I and III.** Pol III is the primary replicase enzyme that performs the elongation of the DNA strand. It adds nucleotides first to the RNA primer and then grows the chain by creating the phosphodiester bonds. It also has a 3'-5' proofreading (exonuclease) function that removes incorrectly incorporated nucleotides. DNA Pol I also has the 5'-3' replicase function, but it is primarily used to fill the gaps in the replicated DNA that occur when the RNA primer is removed. This enzyme also has a 5'-3' exonuclease function that is used to remove the RNA primer.

**Figure 7.** The steps of DNA replication.
5. The DNA replication error is corrected.

6. Meanwhile, Okazaki fragments are synthesized using the lagging strand in a discontinuous manner and leading strand are completed simultaneously.

7. The RNA primers are removed by 5'-3' exonuclease function of DNA polymerase.

8. Replication is completed by the filling in the gaps by DNA polymerae and DNA ligase.

Notes on replication:

**Okazaki fragments**: Both prokaryotic and eukaryotic DNA replication proceed in the 5'-3' direction. This poses a problem because the replication fork on moves in that direction. The problem relates to what is called the *lagging strand*. It must be replicated in a direction that is opposite of the direction of the replication fork. This problem was solved by the discovery of Okazaki fragments (named after the person who discovered the process). In contrast to the *leading strand*, in which DNA is replicated as a single molecule in a *continuous* manner, DNA is replicated in a *discontinuous* manner on the lagging strand. Each of these is primer with a RNA primer, and DNA PolIII in E. coli makes short stretches of DNA. These fragments are then stitched together when the primer is removed and the strands completed by the action of DNA Pol I and ligase.

**Figure 7 (cont.)**. The steps of DNA replication.
Chain Termination Sequencing: the Sanger Technique

A. The Concept

DNA sequencing is the most technique of genomics. By collecting the sequence of genes and genomes we begin to understand the raw material of phenotype development. The most common DNA sequencing is called chain termination sequencing or the Sanger technique (named after the person who created it). It is called chain termination because the incorporation of a dideoxynucleotide terminates the replication process because the nucleotide lacks the required 3'-OH group.

a. A dideoxynucleotide

b. The reaction reagents

- DNA template
- sequencing primer
- dNTPs
- ddNTPs (low concentration)
- DNA polymerase
- salts

c. The sequencing reaction result: fragments that differ by one nucleotide in length

<table>
<thead>
<tr>
<th>Template</th>
<th>Primer</th>
<th>5' A T T C G G A T C C T T A A</th>
</tr>
</thead>
<tbody>
<tr>
<td>5' T A A G C C T A G G A A T T</td>
<td>5' T A A G C C T A G G A A T H 3'</td>
<td></td>
</tr>
<tr>
<td>5' T A A G C C T A G G A A T</td>
<td>5' T A A G C C T A G G A A T H 3'</td>
<td></td>
</tr>
<tr>
<td>5' T A A G C C T A G G A A</td>
<td>5' T A A G C C T A G G A A H 3'</td>
<td></td>
</tr>
<tr>
<td>5' T A A G C C T A G G A</td>
<td>5' T A A G C C T A G G A H 3'</td>
<td></td>
</tr>
<tr>
<td>5' T A A G C C T A G</td>
<td>5' T A A G C C T A G H 3'</td>
<td></td>
</tr>
<tr>
<td>5' T A A G C C T</td>
<td>5' T A A G C C T H 3'</td>
<td></td>
</tr>
<tr>
<td>5' T A A G C</td>
<td>5' T A A G C H 3'</td>
<td></td>
</tr>
<tr>
<td>5' T A A G</td>
<td>5' T A A G H 3'</td>
<td></td>
</tr>
<tr>
<td>5' T A A</td>
<td>5' T A A H 3'</td>
<td></td>
</tr>
<tr>
<td>5' T A</td>
<td>5' T A H 3'</td>
<td></td>
</tr>
<tr>
<td>5' T</td>
<td>5' T H 3'</td>
<td></td>
</tr>
</tbody>
</table>

When a dideoxynucleotide is inserted, the DNA replication process terminates because dideoxynucleotides do not have the necessary free 3' hydroxyl group required for the addition of additional nucleotides. This results in fragments that differ by one nucleotide in length.

Figure 8. The chain termination (Sanger) DNA sequencing technique.
Gel-based Detection of DNA Sequences

A. The concept

Four DNA sequencing reactions are performed. Each contains only one of the four dideoxynucleotides. Each reaction is added to a single lane on the gel. Since one of the dNTPs is radioactive, the gel in which the fragments are separated, can be used to expose an x-ray film and read the sequence.

a. The sequencing products

- **Reaction with ddATP**
  - T T C G G A T C C T T A A
  - 5'  T A A G C C T A G G A A - H 3'
  - 5'  T A A G C C T A G G A - H 3'
  - 5'  T A A G C C T A - H 3'
  - 5'  T A A - H 3'

- **Reaction with ddTTP**
  - T T C G G A T C C T T A A
  - 5'  T A A G C C T A G G A A - T H 3'
  - 5'  T A A G C C T A G G A A - H 3'
  - 5'  T A A G C C T - A H 3'
  - 5'  T A - A H 3'

- **Reaction with ddGTP**
  - T T C G G A T C C T T A A
  - 5'  T A A G C C T A G G - H 3'
  - 5'  T A A G C C T A G - H 3'
  - 5'  T A A G C - H 3'

- **Reaction with ddCTP**
  - T T C G G A T C C T T A A
  - 5'  T A A G C C - H 3'
  - 5'  T A A G C - H 3'

b. The sequencing gel

```
G A T C
3' T
T T
A A
G G
G G
T C
C C
G G
A A
A A
T T
5'
```

The sequencing reactions are separated on a polyacrylamide gel. This gel separates the fragments based on size. The shorter fragments run further, the longer fragments run a shorter distance. This allows the scientists to read the sequence in the 5'-3' direction going from the bottom to the top of the gel.

Figure 9. Gel-based detection of DNA sequencing products.
Fluorescent Sequencing and Laser Detection

A. The Concept

Rather than using four different reactions, each with a single dideoxynucleotide, the advent of fluorescently labeled dideoxynucleotide enabled 1) the sequencing reaction to be performed in a single tube, and the fragment could be detected by laser technology. Originally, the products were separated in a polyacrylamide gel prior to laser detection. The introduction of capillary electrophoresis, coupled with laser detection enabled the detection of up to 96 products at a time.

B. The Reaction Products and Analysis

Sequencing products are loaded on to a capillary electrophoresis unit and separated by size.

Laser detection and software analysis detects the first shortest fragment as ending in a T (thymine). All fragments are detected and interpreted in the same manner.

The Sequence Chromatogram

Figure 10. The fluorescent sequencing and laser detection process of DNA sequencing.