**EDVOTEK**®

**The Biotechnology Education Company ®**

**EDVO-Kit**
# 106

# Principles of DNA Sequencing

**EXPERIMENT OBJECTIVE:**

The objective of this experiment is to develop an understanding of DNA sequencing and analysis. This is a dry lab which contains autoradiographs from an actual DNA sequencing experiment.

**EDVOTEK, Inc. • 1-800-EDVOTEK • www.edvotek.com**

## Table of Contents

## Experiment Components

5    Autoradiographs

## Requirements

A white light box is recommended. The EDVOTEK Visible Light Gel Visualization System (Cat. # 552) is well-suited for this lab. An autoradiograph may also be placed on an overhead projector and shown to the whole class.

## Background Information

The purpose of this exercise is to provide detailed instruction on the practice of DNA sequencing.  The student is provided an actual autoradiograph (an exposed sheet of x-ray film) of a DNA sequencing gel for analysis.  The sequence deduced from the autoradiogram will differ from a wild type sequence by a single nucleotide.  This difference represents an actual mutation in the DNA molecule.  The students are expected to identify the location of the mutated nucleotide.  Students should also work in groups to read the sequence and compare results to assure accuracy.

Rapid analysis of DNA sequence was developed during the 1970's from research groups in the United States and England.  Since its early days, these methods have been refined and automated.

There are two basic approaches to DNA sequence analysis.  One involves a set of organic chemical reactions with the DNA bases.  The other uses an enzymatic process.  The chemical method is tedious and labor-intensive, whereas the enzymatic approach, which is often called the dideoxy method, is quite fast.  The autoradiographs you have been given as part of this experiment, are the result of the enzymatic procedure which uses the Klenow fragment of the *E. coli* DNA polymerase I to make a DNA copy of the region to be sequenced.  More commonly, a DNA polymerase known as sequenase or equivalent is used for DNA sequencing.

A specialized cloning vehicle constructed from an *E. coli* virus, called M13, facilitates rapid DNA sequence analysis. This virus contains a polylinker, which is a short region of DNA, about 57 base pairs, containing several unique restriction sites.  Segments of DNA to be sequenced are inserted into the polylinker region using standard cloning procedures (Figure 1).

The M13 virus contains a single-stranded circular genome with about 7200 nucleotides.  The virus will infect *E. coli* strain JM101 cells which contain a fertility factor.  These cells are F+, and male.  The virus infects by attachment to the sex pilus.  Shortly after infection, the viral DNA will become double-stranded.  It is this form which serves as a template for production of single-stranded DNA progeny.  The DNA associates with the viral proteins to form mature virus and the virus exits the cell by budding; the cell is not lysed.  For cloning and sequencing purposes, double-stranded DNA to be sequenced is inserted into the M13 polylinker region of the  DNA intermediate, and then transformed into competent JM101 *E. coli* cells.  The transformed cells will begin to produce progeny virus.



polylinker region

site of hybridization of sequencing primer

Single-Stranded M13 Viral DNA

**Figure 1**

To sequence DNA which has been inserted into the polylinker region of M13 single-stranded DNA is prepared from viral plaques.  In this experiment, a short 17-base synthetic single-stranded DNA is allowed to hybridize (form a base pairing) with a unique site in M13 adjacent to the polylinker.
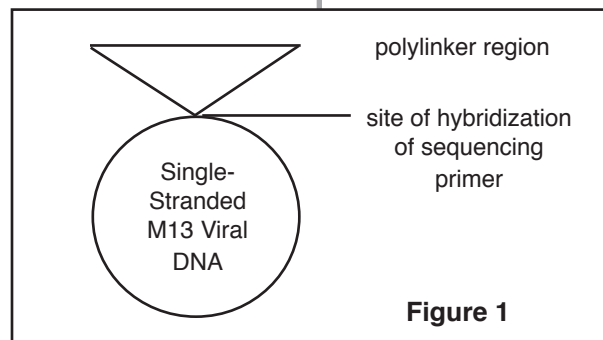
## Background Information



**DNA to be Sequenced**

5'    CTCGGTCGCTCATTACTGGCCGTCGTTTTAC    3'

3'    TGACCGGCAGCAAAATG    5'

**DNA Sequencing Primer**

**Figure 2**



**Deoxynucleotide**

Point of Addition of the Next Nucleotide

**Dideoxynucleotide**

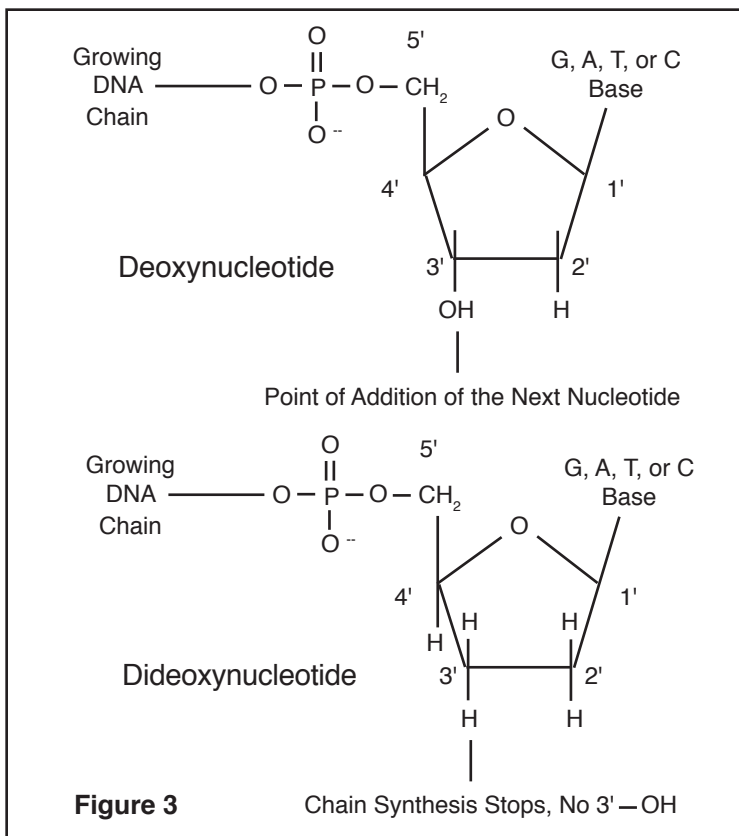**Figure 3**    Chain Synthesis Stops, No 3'—OH

This 17-base oligonucleotide will serve as a primer for DNA synthesis by the Klenow fragment of DNA polymerase I, which lacks the 5'-3' exonuclease activity (Figure 2).

For sequence analysis, four separate enzymatic reactions are performed, one for each nucleotide. Each reaction contains the Klenow fragment of DNA polymerase I, the single-stranded DNA template to which the 17 nucleotide synthetic DNA primer has been hybridized, all four deoxyribonucleotide triphosphates (dATP, dGTP, dCTP, dTTP), $^{32}$P-dATP, and the appropriate buffer for *in vitro* DNA synthesis. In addition, the "G" reaction contains dideoxyGTP, the "C" reaction DideoxyCTP, the "A" reaction DideoxyATP, and the "T" reaction DideoxyTTP. The Dideoxynucleotide concentrations are carefully adjusted so that they are incorporated into a growing DNA strand randomly and infrequently. Once a Dideoxynucleotide is incorporated into a single strand, DNA synthesis is completely terminated (Figure 3). The site of the Dideoxynucleotide incorporation allows one to determine the position of that base. The Dideoxynucleotide lacks a 3'-OH group on the ribose ring and it is impossible for the Klenow fragment of DNA polymerase I to add another nucleotide to the growing strand since a 3'-OH group is absolutely required.

In addition to $^{32}$P-deoxynucleotide triphosphates, isotopic $^{35}$S analogues are also frequently used in DNA sequencing. Non-isotopic methods of using fluorescent dyes and automated DNA sequencing machines are beginning to replace the traditional isotopic methods. In spite of the detection differences, the basic Biochemistry of the dideoxy sequencing method is essentially the same.

Since a particular reaction will contain millions of growing DNA strands, a "nested set" of fragments is obtained; each fragment is terminated at a different position corresponding to the random incorporation of the Dideoxynucleotide.

## Background Information

Figure 4 shows the "nested set" of fragments proclued for a hypothetical sequence in the "G" reaction. The "G" reaction contains dATP, dCTP, dGTP, dTTP, the Klenow fragment of DNA polymerase, an appropriate solution for DNA synthesis, $^{32}$P-labeled dATP and a small amount of dideoxy-GTP.

As can be seen, ddGTP (dideoxyGTP) incorporation randomly and infrequently will produce a "nested set" of fragments which terminate with a ddGTP. The "nested set" is complimentary to the region being sequenced. Similar "nested sets" are produced in the separate "A", "T", and "C" reactions. For example, the "A" "nested set" would terminate with a ddATP.

It should be readily apparent that together the "G, A, T, C" "nested sets" contain radioactive $^{32}$P-labeled fragments ranging in size successively from 19 to 31 nucleotides for the hypothetical sequence in Figure 4.

The "G" reaction contains fragments of 21, 23, 25, 29 and 31 nucleotides in length. The first seventeen of these nucleotides are contained in the synthetic DNA sequencing primer. The rest are added during *de novo* synthesis with the Klenow fragment of DNA polymerase I.

DNA to be Sequenced - Shown in th "G" lane

CTCGGTCGCTCATT

5'                                                    3'

3'              Primer              5'

Fragment Size

ddGTAA _____ 21
          3'              5'

ddGAGTAA _____ 23
          3'              5'

ddGCGAGTAA _____ 25
          3'              5'

ddGCCAGCGAGTAA _____ 29
          3'              5'

ddGAGCCAGCGAGTAA _____ 31
          3'              5'

**Figure 4**

### POLYACRYLAMIDE SEQUENCING GEL

The radioactive reaction products from the G, A, T, and C reactions are applied to separate sample wells on a thin vertical polyacrylamide gel, which is 14 x 17 cm in size. Well # 1 contains the "G" reaction; well # 2 the "A" reaction; well # 3 the "T" reaction; and well # 4 the "C" reaction. The electrophoresis apparatus containing polyacrylamide gel is connected to a power supply with the positive electrode at the bottom and the negative electrode at the top. High voltage is applied (2000 volts, D.C.) to separate the radioactive $^{32}$P-labeled fragments which migrate from top to bottom. The unique sequencing polyacrylamide gels can resolve fragments which differ in size by a single nucleotide, based on size. The smaller fragments move fastest while the larger are slowest.
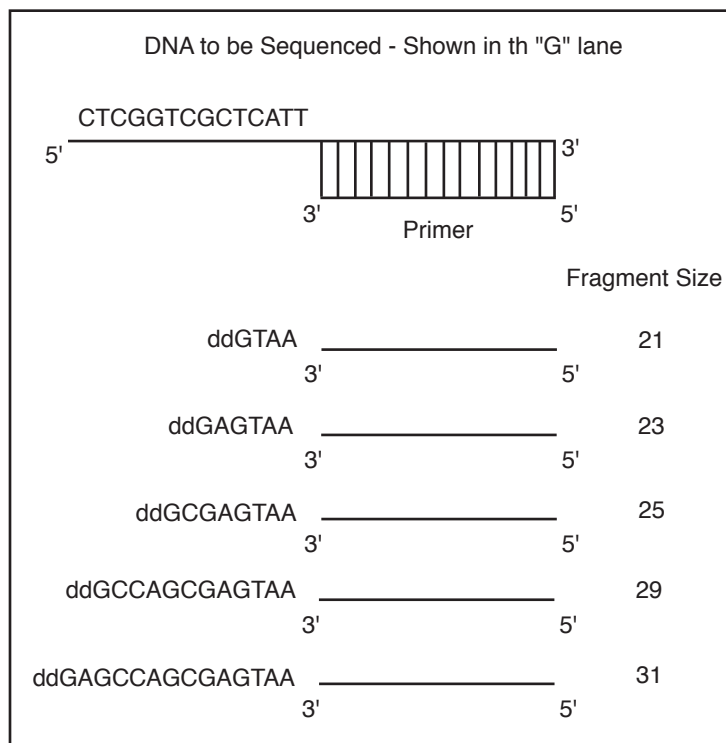
## Background Information

After electrophoretic separation is complete, autoradiography is performed. The poly-acrylamide gel is placed into direct contact with a sheet of x-ray film. Since the DNA fragments are radioactively labeled with 32P, their position can be detected by a dark exposure band on the sheet of x-ray film. For a given sample well, the horizontal "bands" appear in vertical lanes from the top to the bottom of the x-ray film. Generally, a single electrophoretic run can contain 12 sets of "GATC" sequencing reactions since there are 48 sample wells on a typical gel. A single autoradiograph can yield 1000 nucleotides of novel sequence information. Automated sequence machines will yield several multiples of this information.

Figure 5 shows an autoradiograph which would result from analysis of the hypothetical sequence in Figure 4. The dark bands are produced by exposure of the x-ray film with 32P which has been incorporated into the dideoxy-terminated fragments during DNA synthesis. The sequence deduced from the autoradiogram will actually be the complement of the DNA strand contained in the singled-stranded M13 DNA template.
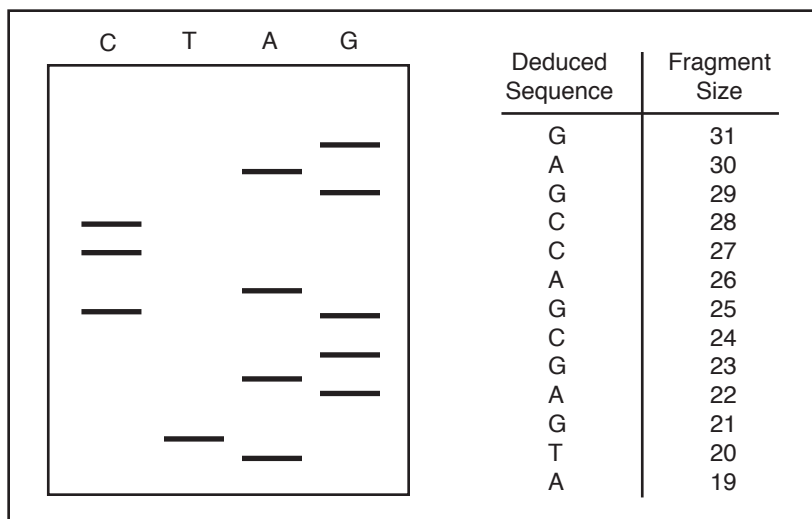
| Deduced Sequence | Fragment Size |
|:---:|:---:|
| G | 31 |
| A | 30 |
| G | 29 |
| C | 28 |
| C | 27 |
| A | 26 |
| G | 25 |
| C | 24 |
| G | 23 |
| A | 22 |
| G | 21 |
| T | 20 |
| A | 19 |

**Figure 5**

## Experiment Overview and Procedures

**EXPERIMENT OBJECTIVE:**

The objective of this experiment is to develop an understanding of DNA sequencing and analysis. This is a dry lab which contains autoradiographs from an actual DNA sequencing experiment.

**EXPERIMENTAL PROCEDURE:**

1. Obtain the sample autoradiograph and place it on a light box to enhance visualization.

2. The sequencing reactions have all been loaded in order: G-A-T-C.

3. Begin analysis of the DNA sequence at the bottom of the autoradiograph with the circled band, which is an A.

4. Compare the deduced sequence to the wild type sequence shown in Figure 6.

5. Identify the location of the mutant nucleotide. What was the mutation? Is there more than one mutation?

> **5'-AGCTTGGCTGCAGGTCGACGGATCCCCGGGAATTCGTAATCATGGT-3'**

**Figure 6**

*Experiment Procedure*

**EDVOTEK**®