Exercise 21
RESTRICTION ENDONUCLEASE DIGESTION OF DNA AND RESTRICTION FRAGMENT LENGTH POLYMORPHISM (RFLP)

Introduction

Present-day DNA technology is partially dependent upon the ability of investigators to cut DNA molecules at specific sites with enzymes called **restriction endonucleases** or **restriction enzymes**. The most useful of these enzymes, which occur naturally within bacteria, recognize particular "target" or **recognition sequences (restriction sites)** within a DNA double helix, and catalyze the cleavage of the DNA (breaking of phosphodiester bonds) at or near those sequences leaving 5'-P and 3'-OH termini. The result of this activity may be the formation of linear segments from closed loops (plasmids) or the formation of discrete DNA fragments of defined length and base sequence. In either case, the ability to selectively cut DNA in this manner is often useful in the manipulation of DNA for scientific purposes.

Among the best-studied restriction endonucleases are those categorized as type II. The activity of type II restriction enzymes in vivo (as they occur naturally within bacterial cells) is coupled with that of **modification enzymes** to form **restriction/modification systems (Type II R-M systems)**. The restriction enzymes enable bacteria to catabolize segments of "foreign" DNA that have entered their cells and provide at least some protection against infection by **bacteriophages**, (since phage DNA is often degraded by their activity). They also probably play a significant role in maintaining species specificity with regard to DNA exchange. Recall that naturally occurring genetic exchange mechanisms function with highest efficiency when the exchange occurs between closely related organisms, and that foreign DNA is usually degraded. The modification enzymes (**methyltransferases**) modify newly formed strands of cellular DNA by adding methyl groups to specific bases such that the cells' own restriction enzymes cannot cut it (i.e., the cellular DNA is protected). With regard to DNA then, these restriction/modification systems allow cells to recognize self versus non-self.

Currently, there are over 3000 different type II restriction endonucleases recognized with 233 different specificities (many of these are commercially available), and a uniform system of nomenclature has been developed to identify them. Since restriction enzymes are found in association with specific bacteria, the names applied to them identify their hosts as follows: The **first letter** of the enzyme name is capitalized, and represents the host **Genus** name, the **next two letters** are not capitalized, and represent the host **species** name (specific epithet). A third letter may or may not be included to indicate the bacterial **strain** or **type** and this may be capitalized or written in lower case. The sequence in which individual enzyme systems were identified is indicated by **roman numerals**, e.g., the first enzyme system identified in *Escherichia coli* strain RY13 is given the name ***Eco*RI**, and the third enzyme system found in association with *Haemophilus influenzae* strain R_d is ***Hin*d III**. Most nucleotide sequences recognized by restriction enzymes are recognized by more than one. The first enzyme system found to have unique specificity (i.e., to recognize a unique nucleotide sequence) is called a **prototype**. Additional enzymes found to have the same specificity as the prototype are called **isoschizomers**. Enzymes recognizing the same nucleotide sequence but cutting in a new location are called **neoschizomers**.

Type II restriction endonucleases recognize and cut DNA strands within recognition sequences that are 4-8 base pairs in length. Many of these have the characteristic of being identical when read 5' to 3' right to left or left to right so are said to be **palindromic**. Others are incomplete or partial palindromes with one or more ambiguous nucleotides situated within their sequences. The cleavage of DNA with such enzymes will give rise to free ends or **termini** that vary in length depending upon where the strands are

cut. Some enzymes yield even cuts, or cuts producing flush- or **blunt-ended** fragments (blunt ends). Other enzymes cut the DNA in a staggered manner giving rise to fragments with protruding 5' or 3' termini 1, 2, 3, or 4 bases in length. These fragments can readily bind (via hydrogen bonding) to other fragments with complimentary base sequences, and so are said to have **cohesive termini** or sticky ends. Enzymes that produce cohesive termini are particularly useful in genetic manipulation procedures because DNA samples from different sources will readily bind together (under the influence of enzymes called **ligases**) to form recombinant strands.

REPRESENTATIVE RESTRICTION ENDONUCLEASES
(with original bacterial sources)

***Hga*I** *Haemophilus gallinarum*	***Taq*I** *Thermus aquaticus*
5'…GACGCNNNNN/N…3' 3'…CTGCGNNNNNNNNNN/N…5'	5'…T/CGA…3' 3'…AGC/T…5'
***Pvu*II** *Proteus vulgaris*	***Hin*dIII** *Haemophilus influenzae*
5'…CAG/CTG…3' 3'…GTC/GAC…5'	5'…A/AGCTT…3' 3'…TTCGA/A…5'
***Alu*I** *Arthrobacter luteus*	***Eco*RI** *Escherichia coli* strain RY13
5'…AG/CT…3' 3'…TC/GA…5'	5'…G/AATTC…3' 3'…CTTAA/G…5'

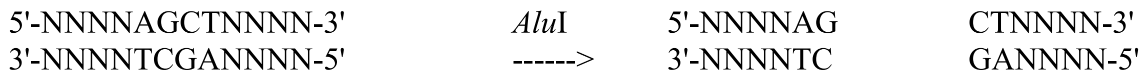
The bacterial plasmids purified during the mini-prep procedure of an earlier exercise can be digested with the restriction enzyme (restriction endonuclease) *Eco*RI to form linear segments capable of binding with other DNA fragments. The enzyme *Eco*RI recognizes and cleaves the DNA base sequence 5'-GAATTC-3' between the G and A residues to generate the four-base cohesive terminus 5'-AATT-3' as shown below.

These plasmids have "**unique**" *Eco*RI cleavage or "cut' sites which means that the enzyme *Eco*RI cleaves the plasmids only once, and will convert these circular molecules into linear forms with cohesive termini at their ends. Any other piece of DNA cut with the same enzyme will have complimentary cohesive termini that can readily bond with the plasmid. This same enzyme cleaves DNA from the bacteriophage lambda in five locations creating six linear fragments, and the *E. coli* genome hundreds of times, generating hundreds of fragments of variable lengths.

When 16S ribosomal DNA is digested by restriction enzymes with short recognition sequences, it is cut in several places to yield multiple fragments (restriction fragments) of different lengths. These

fragments can be subjected to gel electrophoresis and will separate into bands that migrate varying distances from the wells to create a banding pattern known as a **DNA fingerprint** or RFLP pattern. The technique used to generate the fragments and the banding pattern is called **RFLP** for **Restriction Fragment Length Polymorphism**. Each RFLP band is actually an accumulation of DNA fragments that are all the same length. Even though the nucleotide sequences of 16S ribosomal DNA are highly conserved, this DNA obtained from bacteria of different species will not be identical. There is generally about 1% variation in the base sequence (16 bases out of 1600) between bacteria categorized into different species. This variation in base sequence will influence the number and position of restriction enzyme recognition sequences, and therefore the number and position of cut sites. When 16S ribosomal DNA from different bacterial species is amplified with the PCR and then cut with the same restriction endonuclease, the RFLP patterns formed will be different, i.e., the number and size of their respective restriction fragments will be different. For this reason restriction endonucleases, and RFLP can be used in bacterial identification.

In this exercise you will use the restriction endonuclease ***Alu*I** (originally obtained from *Arthrobacter luteus*) to digest the bacterial DNA amplified with the PCR during an earlier exercise. This enzyme recognizes and binds with the base sequence 5'-AGCT-3' and cuts between the guanine and cytosine to yield **blunt ends** as shown below.



Because the recognition sequence for *Alu*I is short, the enzyme can be expected to bind with and cut 16S ribosomal DNA in five or more places, generating unique RFLP patterns. DNA digested prior to this exercise will be subjected to gel electrophoresis, and analysis of the resulting RFLP patterns will provide information applicable to the identification of Physiological Unknown #2.

Procedure

1. Obtain a tube containing digested DNA for your Physiological Unknown #2 culture from the numbered set provided. Quickly thaw the contents and place the tube in the ice bucket provided.

2. Follow the procedural steps described in Exercise 20 – **Gel Electrophoresis of DNA Samples**, Procedure (Exercise B).

3. When you have loaded your DNA sample into a gel, obtain a PowerBook from the mobile cart and access the Microbiology web site.

4. Go to web-based laboratory assignments and access the grid containing **Physiological Unknown #2 (PUNK2) – Word Files** (verify that the semester and year are correct).

5. Click on your number in the grid to access the Word File containing the most complete version 16S rDNA nucleotide sequence available for your unknown.

6. Using the "Find" option of Microsoft word (Apple F), fine all of the recognition sequences AGCT within your 16S ribosomal DNA sequence and convert them to AG….CT. This can be accomplished as follows. Highlight the sequence, move the cursor to "edit" upper right, click and hold while scrolling down to "Find", and then release. When the "Find" window opens, type AGCT in the space to the right of "Find what:". Click on the replace tab and type "AG….CT" into the space to the right of "Replace with:". Then click on "Replace all".

Note – AGCT is the **recognition sequence** for the restriction endonuclease AluI used to cut your PCR product DNA. The enzyme binds and cuts DNA between the Guanine and Cytosine bases. When DNA is cut, the resulting fragments are called restriction fragments or **RFLP**s.

7. Look for the AG….CT sites in your sequence and record the number of times AluI will cut your DNA.

8. Determine how long each restriction fragment will be by using the "word count" option of Microsoft Word. This can be accomplished as follows. Highlight each restriction fragment (one at a time), move the cursor to "Tools" middle of tool bar, click and hold while scrolling down to "word count", and then release. The word count window will tell you how many characters are in each selection, i.e., how many bases are in each restriction fragment. **Record this information**. It will be applied to the identification of your Physiological Unknown #2.

9. Make a diagrammatic representation of an electrophoresis gel containing the RFLP pattern generated by cutting the 16S r-DNA from your PUNK2 with AluI. Remember that smaller fragments will travel faster and farther in a gel than will larger ones.

10. Compare your RFLP diagram with the RFLP generated in the electrophoresis gel. Use the 1Kb ladder to estimate the sizes of the restriction fragments visible in your lane of the electrophoresis gel and compare these to the sizes of fragments generated during the computer-simulated digest. Consider what factors might cause the patterns generated to be different.

Questions

1. What are restriction endonucleases? Where are they normally found?

2. What function do restriction endonucleases have in vivo?

3. What name would be given to the second restriction endonuclease system found in association with *Bacillus stearothermophilus* strain ET?

4. What function do restriction enzymes have in recombinant DNA technologies?

5. What are the differences between blunt and cohesive termini? Which of these is produced when DNA is cut with EcoRI and which when DNA is cut with AluI?

6. What is RFLP, and what role does this technique play in bacterial identification?

Exercise 21 - Supplement
CHARACTERIZATION OF RESTRICTION FRAGMENTS

Introduction

When a restriction enzyme such *Eco*RI is used to cut chromosomal DNA, the result is the formation of hundreds of DNA fragments of variable length. Once generated, these fragments can be inserted into plasmids or other **cloning vectors** using DNA **ligase**, an enzyme that catalyzes the formation of phosphodiester bonds between adjacent nucleotides. A DNA fragment that has been inserted into a plasmid is said to have been molecularly **cloned** away from other DNA fragments, and the plasmid-chromosome combination is referred to as **recombinant DNA**. In a typical cloning experiment, a large number of DNA fragments are generated, inserted into cloning vectors (cloned) and then reproduced within bacterial cells. A collection of such fragments is referred to as a **gene bank** or **gene library**.

The probability that a particular gene will be found within a gene library is influenced by the size of the gene, the average size of the DNA fragments cloned, and the total number of recombinant molecules produced. The greater the number of DNA fragments cloned, and the larger these fragments, the greater the probability that every gene within the chromosome will be present within at least one of the recombinant molecules formed.

Restriction fragments may be **characterized**, i.e. their size determined, through electrophoresis. Fragments that have been cloned in plasmids can be removed from their vectors via restriction endonuclease digestion, and will form separate bands within an agarose gel. If the DNA bands representing the various fragments are compared to the bands of a known standard, it is possible to determine their relative sizes. It is also possible to determine the sizes of fragments within plasmids (without digestion) by determining the size of the plasmid-restriction fragment combination, and then subtracting the size of the plasmid.

During this exercise, fragments of DNA cloned from the chromosome of *E. coli* and carried within the cloning vector pUC19 will be characterized. These fragments were cloned by cleaving chromosomal DNA with the restriction enzyme *Eco*RI, and ligating the resulting fragments into pUC19. The plasmids were used to transform bacteria and then were replicated many times over within actively growing bacteria cells. Since the plasmid pUC19 has only one *Eco*RI cleavage site (a unique *Eco*RI site), restriction fragments can easily be removed from this plasmid via digestion with the same enzyme. For this exercise, digested and undigested DNA samples were subjected to electrophoresis along with DNA from the bacteriophage lambda. Lambda is used as a standard for calibration because the nucleotide sequence of the lambda genome is known, along with the cleavage sites for various enzymes; and digestion of lambda DNA with specific enzymes will usually yield a predetermined number of fragments each of a specified size.

Procedure:

1. Obtain a pre-made agarose gel, place it correctly in a gel box (electrophoresis chamber) containing an appropriate quantity of 1X TBE buffer.

2. Obtain samples of the DNA provided (recombinant DNA and Lambda standard DNA), a sample of loading dye, a strip of parafilm and a 0.5-10μL pipette. Place the DNA samples on ice.

3. Using a different clean, sterile pipette tip for each sample, mix 5μL of each plasmid DNA sample with 5μL of dye on the parafilm, and then load each total volume (10μL) into a different well in the gel. Place lambda standard DNA in the two outside wells, and be sure to record the order of the other four samples.

4. Place the lid on the box, turn the power supply on, adjust the voltage applied to the electrodes to about 45 volts (flat bottom box) or 130 volts (deep well chamber), and allow electrophoresis to proceed until the color band produced by the Bromophenol-blue moves to a point about three quarters of the way to the end of the gel, and then turn the power off.

.

5. Ask the instructor to stain the gel with ethidium bromide or GelRed and place it on the transilluminator for observation. If the gel is suitable (DNA bands clearly visible) obtain a photograph of the preparation or lay a ruler beside the gel and quickly sketch the relative positions of the bands observed.

6. Using the Lambda DNA fragments as molecular weight standards, determine the relative sizes (molecular weights) of the restriction fragments present. Remember, you must subtract the size of the vector from the total to determine fragment size.

Questions:

1. What is a cloning vector? What is a gene library?

2. What were the sizes of the restriction fragments present within your DNA samples?

3. If you were looking for a gene that contained 1500 base pairs (1.5 Kb), could this gene be present in any of the recombinant molecules you were working with?